GENETIC DIVERSITY IN THE *LEISHMANIA DONOVANI* COMPLEX

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy of the University of London

August 2000

Isabel Larguinho Mauricio

Pathogen Molecular Biology and Biochemistry Unit

Department of Infectious and Tropical Diseases

London School of Hygiene and Tropical Medicine

University of London

United Kingdom

Abstract

The *Leishmania donovani* complex comprises four described species: *L. donovani, L. archibaldi, L. infantum* and *L. chagasi*. *L. chagasi* is the only New World species and has been considered similar to *L. infantum*, although some authors insist on maintenance of its independent species status. *L. donovani* has at least two major epidemiological subgroups whose relationships are poorly understood.

In this thesis, molecular biological techniques were used to investigate the taxonomy and phylogenetic relationships within the *L. donovani* complex, with isoenzyme analysis (IEA) as reference technique. Random amplification of polymorphic DNA (RAPD) was used to provide anonymous genetic markers which allowed overall comparisons of genomes. Selected target genes and intergenic regions were also amplified by the polymerase chain reaction (PCR), namely the major surface protease (msp or gp63), the mini-exon and the ribosomal internal transcribed spacer (ITS). PCR products of intergenic regions between *msp* genes (ITG/CS and ITG/L), mini-exon and ITS were analysed by restriction fragment length polymorphism (RFLP). Phylogenies generated from each of the methods were compared with that of IEA.

L. infantum and *L. chagasi* were found to be synonymous, whilst *L. donovani* was found to be more polymorphic than *L. infantum* and a fourth possible species in the complex, *L. archibaldi*, was not supported. Six genetic groups of strains were identified in the *L. donovani* complex, based on all DNA based analyses, which agreed with IEA typing. Pooled data from RFLP and RAPD analyses generated robust phylogenies which were congruent with ITG/CS RFLP and *msp* DNA sequence based phylogenies, but not with IEA phylogenies. The evolutionary history of the *L. donovani* complex is analysed in the light of the present results. The diverse typing methods were also evaluated and genetic markers suggested, that are applicable to classification and typing of *L. donovani* species and strains.

Contents

List of figures

List of tables

LIST OF ABBREVIATIONS

6PGD - 6-phosphogluconate dehydrogenase (decarboxylating)

ACL - American cutaneous leishmaniasis

AFLP - amplified fragment length polymorphism

AIDS - acquired immunodeficiency syndrome

ALAT - alanine aminotransferase

AP-PCR - arbitrarily primed PCR

ARI - *Phlebotomus ariasi*

ARV - *Arvicanthis* sp.

ASAT - aspartate aminotransferase

ASD - average percent standard deviation

bp - base pair

BR - Brazil

CAN - *Canis familiaris*

CC - cophenetic correlation coefficient

CD1, 2, 3, ... - cluster of differentiation 1, 2, 3, ...

CER - *Cerdocyon thous*

CL - cutaneous leishmaniasis

CLB - cell lysis buffer

CN - China

CR3 - complement receptor 3

CY - Cyprus

DAT - direct agglutination test

DAS - *Dasypus novemcinctus*

DCL - diffuse cutaneous leishmaniasis

ddN(A/C/G/T)TP - dideoxy (adenine/cytosine/guanine/tymine) triphosphate

DGGE - denaturing gradient gel electrophoresis

DIA - NADH diaphorase

EDTA - ethylenediaminetetraacetic acid

ELISA - enzyme-linked immunosorbent assay

ET - Ethiopia

EtBr - ethidium bromide

FR - France

G6PD - glucose-6-phosphate dehydrogenase

GLUD - glutamate dehydrogenase

GOT - glutamate-oxaloacetate transaminase

gp63 - glycoprotein 63

GPI - glycosyl-phosphatidylinositol or glucose phosphate isomerase

HEPES - N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulforic] acid, sodium salt

HIV - human immunodeficiency virus

HN - Honduras

HOM - *Homo sapiens*

ICD - isocitrate dehydrogenase

IEA - isoenzyme electrophoresis analysis

IEF - isoelectric focusing

IFAT - immunofluorescence

IFNγ - interferon γ

IHMT - Instituto de Higiene e Medicina Tropical

IL - interleukin

IN - India

iNOs - inducible nitric oxide synthase

IQ - Iraq

IR - Iran

IT - Italy

ITG - intergenic region

ITG/CS - intergenic region between *mspC* and *mspS4* genes

ITG/L - intergenic region between *mspL* genes

ITS - internal transcribed spacer

kb - kilobase pairs

KE - Kenya

LB - Lebanon

LDH - lactate dehydrogenase

LSHTM - London School of Hygiene and Tropical Medicine

LSU - large sub-unit

LUT - *Lutzomyia* sp.

MAR - *Phlebotomus martini*

MCL - mucocutaneous leishmaniasis

MDH - malate dehydrogenase

ME - malic enzyme

MEM - minimum essential medium

MER - *Meriones* sp.

MHC - major histocompatibility complex

min - minutes

MPI - mannose phosphate isomerase

mRNA - messenger RNA

msp - major surface protease

mspC - constitutive major surface protease gene

mspC3 - 3' end of the constitutive major surface protease gene

mspL - logarithmic phase major surface protease gene

mspS - stationary phase major surface protease gene

MT - Malta

MTT - 3-94,5-dimethylthiazol-2-yl0-2,5-diphenyl-2 H-tetrazolium bromide

My - million years

NAD - β-nicotinamide adenine dinucleotide

NADH - β-nicotinamide adenine dinucleotide, reduced form

NADP - β-nicotinamide adenine dinucleotide phosphate

NH - nucleoside hydrolase

NK - natural killer

NNN - Novy, McNeal and Nicolle agar medium

NO - nitric oxide

NP - purine nucleoside phosphorylase

nt - nucleotide

NTS - non-transcribed spacer

NW - New World

OD - optical density

OW - Old World

PA - Panama

PBS - phosphate buffer saline

PBSS - proline balanced salts solution

PCR - polymerase chain reaction

PE - Peru

PEPD - prolime iminopeptidase

PFGE - pulse-field gradient gel electrophoresis

PGM - phosphoglucomutase

PKDL - post-kala-azar dermal leishmaniasis

PMS - phenazine methosulphate

PT - Portugal

RAPD - random amplified (amplification of) polymorphic DNA

RFLP - restriction fragment length polymorphism

rRNA - ribosomal RNA

RS - reference strains

S-α-MEM - suplemented α-Minimum Essential Medium

SA - Saudi Arabia

SD - Sudan

SDS - sodium dodecyl sulphate

SER - *Sergentomyia minuta*

s.l. - sensu lato

SOD - superoxide dismutase

SP - Spain

spp - species

SQ - sum of squares

s.s. - sensu stricto

SSC - sodium citrate buffer

SSCP - single stranded conformational polymorphism

SSU - small sub-unit

SU - former Soviet Union

TAE - Tris acetate EDTA buffer

TBE - Tris-borate-EDTA buffer

TEMED - NNN'N' tetramethylethylenediamine

TGFβ - tumour growth factor β

Th - T helper cell

TNFα - tumour necrosis factor

Tris - Tris(hydroxymethyl)aminomethane

TU - Tunisia

UPGMA - unweighted pair-group method with arithmetic mean

USA - United States of America

UV - ultra-violet

VL - visceral leishmaniasis

VUL - *Vulpes vulpes*

v/v - volume / volume

w/v - weight / volume

WHO - World Health Organization

Acknowledgements

I am very grateful to Professor Pedro Abranches and Dr. Virgilio do Rosario, and colleagues and friends Gabriela Santos-Gomes, Lenea Campino for encouraging me to look further ahead and support my scientific interests. My friend Sandra Pereira for providing my first accomodation in London.

To Praxis XXI for personal financial support.

I will always be in debt to Professor Michael A. Miles for believing in me from the start and for his personal and scientific support.

My laboratory colleagues and friends Debbie Nolder, Michael Gaunt, Russell Stothard, Iain Frame and Hernan Carrasco. Friends in the laboratory who shared problems and good times, Matthew , Nahla, Sueli, Patricia, Dia, Suzana.

The isoenzyme and *Leishmania* experts David Evans and Leila Rustamani.

My upgrading examiners David Conway and Douglas Barker, for useful criticisms and suggestions.

My 'old' office mates and friends Aura, Angeles, Liz, Samad and all that passed through our room. For useful discussions (not always scientific...) and moral support.

I wish to thank all those whom I forgot to mention, but were there for me.

I really appreciated the British weather which kept me in the lab most of the time, especially weekends and holidays.

I am indebted to my family, especially my parents, who always supported me in the pursuit of academic excellence and happiness.

Last but not least, my loved, now husband, Luis who was SO patient with crazy timetables and strange humour, and, most of all, whose support kept me sane through the most difficult times.

Finally, I dedicate this thesis to my parents, brother, sister and to the love of my life, Luis.

1. Introduction

1.1. Leishmaniasis

1.1.1. Characterization of leishmaniasis

1.1.1.1. One name for a wide range of disease forms

Leishmaniasis is a disease which results from infection with parasites of the genus *Leishmania*, which are transmitted through the bite of sandflies, diptera (Fam Psichodidae) of the genus *Phlebotomus* in the Old World (OW) or *Lutzomyia* in the New World (NW). The disease has a tropical to sub-tropical distribution, which is mostly dependent on the distribution of the vectors.

All continents, except Australasia and Antarctica, have endemic areas for leishmaniasis, which is mainly present in Southern Europe, North and Central Africa, the Near and Middle East, the East of China and Central and South America, in a total of 88 countries (http://www.who.int/inf-fs/en/fact116.html. Leishmaniasis affects at least 12 million people, with 2 million estimated new cases each year (http://www.who.int/inf-fs/en/fact116.html), of which 1.5 million are new cases of cutaneous leishmaniasis (CL) and 0.5 million are new cases of visceral leishmaniasis (VL), but these figures are probably underestimated (Desjeux, 1992). The most affected areas are in developing countries or include impoverished populations. Three hundred and fifty million people are potentially at risk, especially in rural and poorly developed urban areas. Massive migrations to or from endemic areas (de Beer *et al.*, 1991; Mengesha and Abuhoy, 1978), often accompanied by famine, have been followed by epidemics but invasion of forest areas seems to be a further risk factor in tropical countries (Molyneux, 1998).

Human leishmaniasis includes a variety of clinical presentations - visceral (VL) or kala-azar, mucocutaneous (MCL), cutaneous (CL), diffuse cutaneous (DCL) and post kala-azar dermal leishmaniasis (PKDL) - which usually correlates with the causative species (Table I). Visceral leishmaniasis is caused by the *L. donovani* complex (including *L. infantum* and *L. chagasi*). *Leishmania donovani* is also the aetiological agent of PKDL. *Leishmania major* and *L. tropica* cause Old World CL whilst *L. braziliensis* complex parasites cause MCL and the *L. mexicana* complex causes American cutaneous leishmaniasis (ACL) and also DCL.

The relationship between the parasite species and the type of leishmaniasis is not always straightforward. *Leishmania donovani* complex strains are known to cause CL (Ben Ismail *et al.*, 1992; del Giudice *et al.*, 1998; Mebrahtu *et al.*, 1993; Oliveira Neto *et al.*, 1986a; Ponce *et al.*, 1991) or even MCL (Alvar *et al.*, 1990; el Hassan *et al.*, 1995), and the diverse New World species cause several types of often

indistinguishable cutaneous - mucocutaneous leishmaniasis. *Leishmania tropica* has also been found to be agent of mucocutaneous lesions in the Old World (Mohammed and Wright, 1987; Morsy *et al.*, 1995), although not as severe as in the NW, and to cause VL in American troops during the Gulf War in the Persian Gulf between 1990 and 1991 (Hyams *et al.*, 1995). Host susceptibility, probably dependent on genetic factors (Bryceson, 1996) and immune or nutritional status (Cerf *et al.*, 1987; Dye and Williams, 1993), seems to play an important role on development of disease. External factors, such as the inocula, the vector saliva and the parasite genotype, however, may also be important.

Since treatment given to each patient is dependent on the parasite and on early diagnosis, it is important to identify the responsible agent correctly and quickly. It is also important to assess the type and intensity of the patient's immune response to try and predict the outcome of an initial infection.

Table I - *Leishmania* species and disease forms

Type of leishmaniasis	Main species	Other species
Visceral	*L. donovani* complex	-
Old World cutaneous	*L. major*, *L. tropica*	*L. donovani* complex
Mucocutaneous	*L. braziliensis* complex	*L. tropica*
American cutaneous	*L. mexicana* complex	*L. braziliensis* complex
Diffuse cutaneous	*L. mexicana* complex	-
Post kala-azar dermal	*L. donovani*	-

1.1.1.2. Clinical presentation

Visceral leishmaniasis is the most serious form of leishmaniasis. The fatality rate for symptomatic cases is high without treatment (up to 90%) (Seaman *et al.*, 1996; Thakur, 1984), especially in impoverished populations with deficient nutrition. With treatment and in healthy populations VL is much less serious (al Jurayyan *et al.*, 1992; Evans *et al.*, 1995; Jeronimo *et al.*, 1994; Saxena *et al.*, 1996), although there is significant morbidity (Siddig *et al.*, 1990). The *Leishmania* infected macrophages concentrate in the liver, spleen bone marrow and lymph nodes, thus causing hepatomegaly, splenomegaly, anaemia and lymphadenopathy. VL is also characterised by intermittent fever and weight loss. There is depletion of cell-mediated immunity but humoral response is very strong. The darker coloration of the skin in the Indian form of disease led to the Sanskrit name of kala-azar, for black fever.

Most CL cases are single, painless cutaneous lesions which are self-healing or easy to treat, although disfiguring. The appearance of the lesions may vary enormously, from dry to wet sores and from single to multiple lesions. All *Leishmania* species isolated from humans have been implicated as aetiologic agents of CL, although those mainly responsible are *L. major* and *L. tropica* in the OW (Ashford and Bates, 1999) and *L. mexicana*, *L. guyanensis* and *L. braziliensis* complexes in the NW (Lainson and Shaw, 1999). Simple CL can be the primary lesion before the development of MCL or DCL.

Mucocutaneous leishmaniasis involves destruction of the mucocutaneous border and cartilage and can be fatal if the respiratory mucosa of the larynx and pharynx are affected. It is thought that parasites metastise to the facial cartilage and mucosa from primary infections. Proximity of the primary lesion to the face and high number of sandfly bites seem to correlate with development of MCL. The highly disfiguring and often debilitating lesions on the face, may lead to social stigmatisation of the patient. Diagnosis and identification of the parasite is not easy since isolation of parasites is not often achieved. Treatment is very difficult and lesions often require intense reconstructive surgery.

Diffuse CL results from skin restricted metastasis, usually by *L. mexicana* complex parasites in the NW (Lainson and Shaw, 1999) and in the OW by *L. aethiopica* (Ashford and Bates, 1999), producing nodular lesions throughout the body. The condition is characterised by anergy of both cell-mediated and humoral immunity and is extremely resistant to therapy. In the Dominican Republic an unknown species of *Leishmania* was also found in DCL type lesions (Lainson and Shaw, 1999).

In *L. donovani* endemic areas some patients develop skin lesions, a condition known as PKDL (Ashford and Bates, 1999) because it usually follows therapeutic 'cure' from VL (from before cure of VL up to at least 10 years after). There are cases, however, without any previously known history of VL, although sub-clinical infection might have been present. Clinical presentation is very variable, from alteration of pigmentation to cutaneous nodules, mainly in surfaces exposed to light, and the condition can become equivalent to DCL.

1.1.1.3. History

Although the aetiological agents of leishmaniasis only began to be identified at the beginning of the 20th century, the presence of the disease in the human population may be quite ancient. Pottery in the NW depicts faces with mutilations very similar to those seen today in MCL cases (Ashford and Bates, 1999), although

some authors argue that they may be ethnic mutilations (Jarry, 1999). Similarly, there are reports of afflictions similar to OW CL from 5000 years ago in what is now Iraq. In the 10th century there is an indisputable description of CL by an Arab physician (Jarry, 1999). The first description of VL is much more recent (1835), from Greece, and literature from the Indian continent did not describe any cases until 1872, when serious epidemics at the beginning of that century were reported (Jarry, 1999). Although visceral affliction is prone to being mistaken with malaria or plague, it is possible that Indian VL was a recent introduction from areas with low endemicity, such as Northeast Africa.

1.1.1.4. Distribution

Visceral leishmaniasis can be roughly divided into three epidemiological types. Human restricted VL in the Indian sub-continent is caused by *L. donovani* (*s.s.*); the canine zoonosis, which afflicts mainly children in the Mediterranean basin through the Asian interior to Meridional China and the Americas, including countries from Brazil to the United States of America (USA) caused by *L. infantum* and *L. chagasi*; and the form in North-eastern Africa, without an identified reservoir, caused by *L. donovani* (*s.l.*).

Old World CL due only to *L. major* is present in North Africa and the borders with the South Sahara desert, due to either *L. major* or *L. tropica* from the Arabian peninsula to West of the Indian Continent, and due only to *L. aethiopica* in Ethiopia.

In the New World (Lainson and Shaw, 1999), of the species of the *L. mexicana* complex; *L. mexicana* is present in the USA, Mexico, Belize, Guatemala, Honduras and Costa Rica, *L. amazonensis* in Bolivia, Brazil, Colombia, French Guyana and Paraguay. Of the species of the *L. braziliensis* complex, *L. braziliensis* is reportedly present in most Latin American countries except perhaps Argentina, although it may often have been inadequately identified, and *L. peruviana* is in Andean Peru and perhaps other Andean countries. Of the species in the *L. guyanensis* complex, *L. guyanensis* is present in Brazil, north of the Amazon, and the Guyanas but also in Equador, Venezuela and Peru, *L. panamensis* is present in Panama but also in Colombia, Equador, Venezuela, Costa Rica, Honduras and Nicaragua.

1.1.2. Visceral leishmaniasis epidemiology - The Portuguese example

Leishmaniasis caused by *L. infantum* is considered sporadic in continental Portugal, with a number of endemic regions. Currently identified endemic regions are Vila Real (North), the Algarve, Greater Lisbon and Évora (Figure 1). The Algarve (Campino *et al.*, 1995) and Évora (Semião Santos *et al.*, 1995) have only been fully recognised as endemic areas in the last five years. According to data from the Primary Health Care General Direction (Direcção Geral dos Cuidados Primários de Saúde) in 1991, the annual incidence was of 8.3 / 100 000 inhabitants in the Alto Douro, 0.2 / 100 000 inhabitants in the Lisbon Metropolitan Region and 1.2 / 100 000 inhabitants in the Algarve.

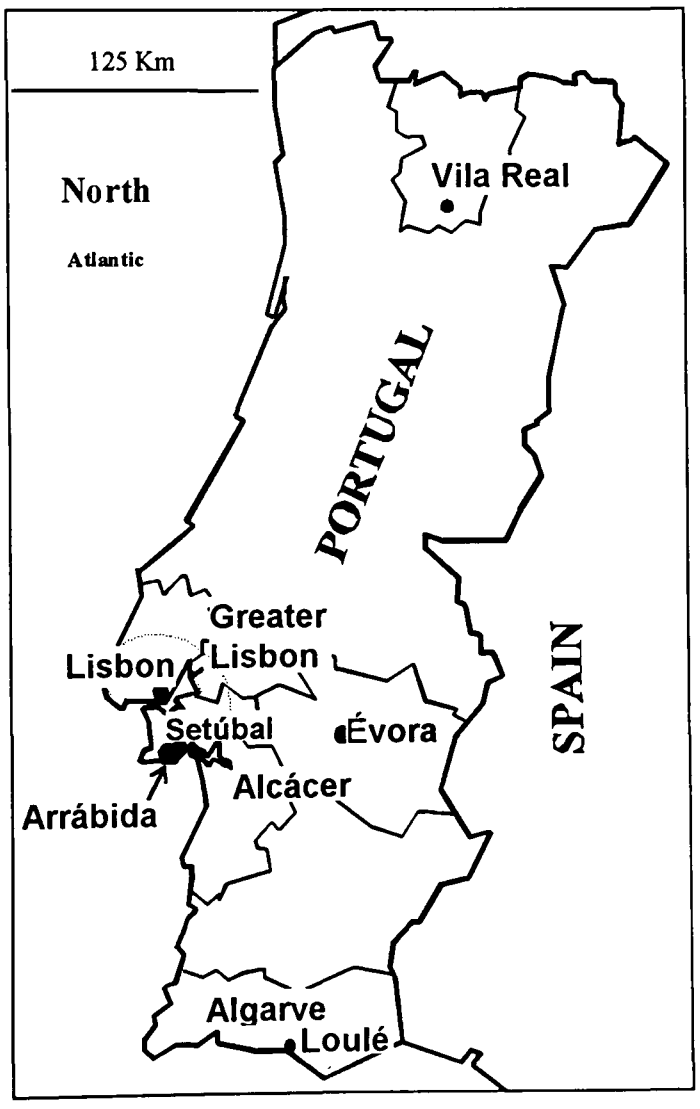


Figure 1 - Distribution of foci in Portugal

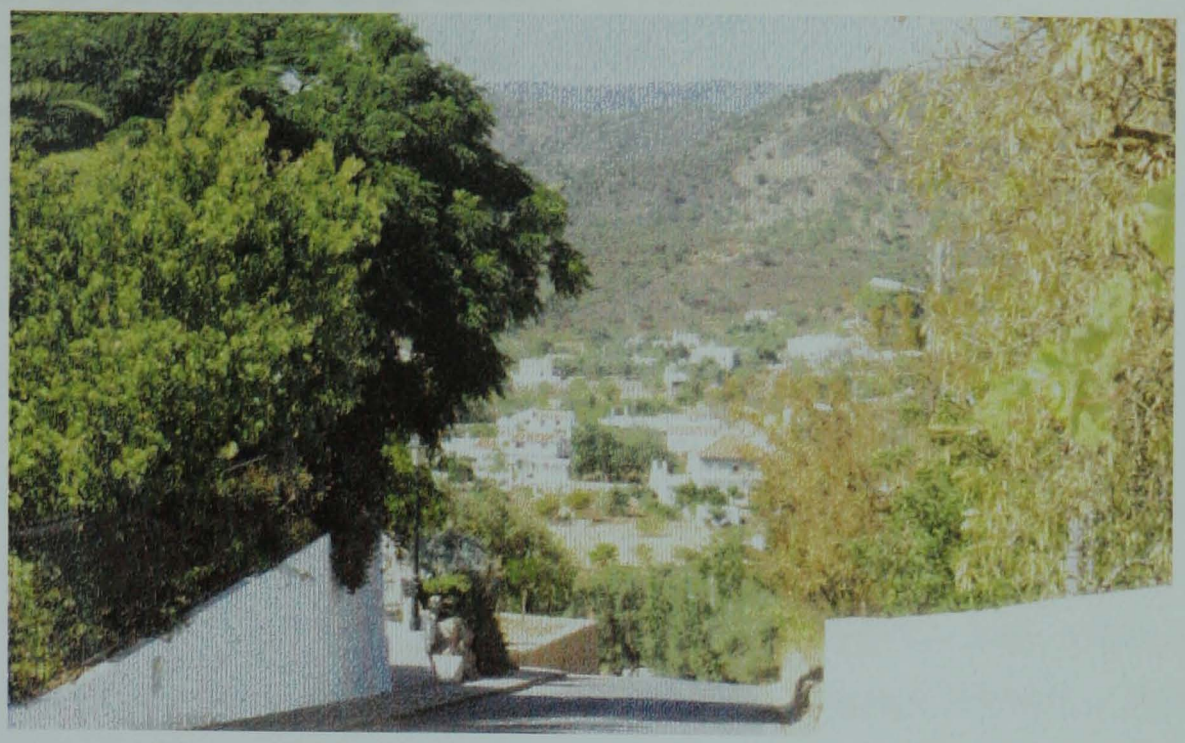
Prevalence of *Leishmania* infection in dogs was found to be of 11.4% in Vila Real (Figure 2) (Abranches *et al.*, 1992; Sampaio-Silva *et al.*, 1993), with a maximum of 37.8% in a locality called Vale de Mendiz (Abranches *et al.*, 1992). In Greater Lisbon region, prevalence varied from 8.8% to 3.8 % in rural (Figure 3) and urban areas, respectively (Abranches *et al.*, 1987). A maximum prevalence of 10% was recorded in the Natural Park of Arrábida (Figure 3), where a sylvatic cycle was identified with the red fox *Vulpes vulpes* as reservoir (prevalence of 5.6%). In the

region of Lisbon the human prevalence correlated negatively with the canine prevalence. In the Algarve, an overall prevalence of 7% was found in the Loulé county, with a maximum of 18.4% in the locality of Querença (Figure 4). In the Évora district an overall canine seroprevalence of 3.9% was found with a maximum of 6.9% in the locality of Nossa Senhora dos Aflitos.

Figure 2 - North of Portugal, Alto Douro - Alijó - vegetation and landscape.

Figure 3 - Arrábida - view of the Serra and adjacent rural areas

A



B

Figure 4 - Algarve: Querença. A - Village and surroundings. Vegetation and landscape in the 'barrocal': rural hilly area with high population density. B - Kennel: dogs roam freely or are restricted to unprotected kennels.

Epidemiology of visceral leishmaniasis is poorly understood, and Portugal is not an exception. Four main foci have been identified, but only a few counties have been studied, and many regions may have been overlooked. In the focus of Alcácer do Sal (Figure 5), in which no reservoir cases were found by Abranches *et al.* (1983), at least 10 cases of canine visceral leishmaniasis were reported later (Galrito *et al.*, 1991; Vicente, 1990). Given that high incidence of canine leishmaniasis may not correlate with high incidence of human leishmaniasis (Abranches *et al.*, 1987), and

that not all seropositive dogs develop symptoms, the prevalence of *Leishmania* infection in regions not considered endemic may be high enough to maintain endemicity throughout the country (Rioux *et al.*, 1972). Phlebotomine sandflies prefer areas with moderate temperatures (Alves Pires and Ribeiro, 1991), but are fairly tolerant to drought. A nation-wide epidemiological enquiry would be required in order to know the real situation of Portugal regarding leishmaniasis. Prevalence rates among stray dogs and sylvatic canids are also unknown, and very difficult to study. It is thus not known if leishmaniasis is endemic in the entire territory or if there are true limited foci. It must be noted that in the studied foci there were high prevalence regions in the borders (Abranches *et al.*, 1992; Campino *et al.*, 1995; Sampaio-Silva *et al.*, 1993; Semião Santos *et al.*, 1995), which may indicate that each focus is larger than the studied area. Since effective field study of *Leishmania* infection in Portugal is so difficult, another way to assess whether the foci are limited or if there is a continuum across the country is to study the genetics of the parasites themselves.

Figure 5 - Alcácer do Sal, a humid rice growing region with a history of malaria, which was eradicated by insecticide spraying, co-existing with more dry areas with *Quercus* sp. Previously described as an endemic focus, leishmaniasis was apparently eradicated following anti-malaria campaigns. The resilience of leishmaniasis could be explained by the higher resistance to drought by *Phlebotomus* or by re-introduction from nearby endemic areas, such as the Arrábida.

Until recently, the only *Leishmania* zymodeme identified in Portugal had been MON 1. In the last 20 years, however, two MON 24 isolates were found infecting the sandfly vector in the endemic area of Vila Real and in an human immunodeficiency virus (HIV) co-infection case. A case of CL caused by MON 29 was also found near the border with Spain in the Alentejo region of Portugal. It was suggested that it could

be due to importation of Spanish strains to the Portuguese territory. However, the Alentejo has not been studied, except for the municipality of Évora, although reports from the decades of 60 and 70 indicated high incidences. It is possible that there is a continuum with Spain with presence of more unusual zymodemes in border areas. An HIV patient was also found with *L. donovani* MON 18 (Campino *et al.*, 1994), probably a needle transmission case. There are unreported cases of unusual zymodemes or *Leishmania* like parasites found among HIV patients in Portugal; but to a lesser extent than in Spain.

Typically a disease of young children in Portugal, VL is becoming more and more a disease of the immunologically compromised, as in other European and American countries. Most affected adults are co-infected with HIV, although other immunocompromised patients, notably transplant recipients, are also at risk. Most HIV / *Leishmania* co-infected individuals are men and drug users (Campino *et al.*, 1997). As in other countries, parasites have also been found in unusual organs, such as the lungs, intestines and skin, which may be a consequence of *Leishmania* being allowed to spread uncontrolled. However, it may be that lesions are revealed by the intensive screening for opportunistic diseases done on acquired immunodeficiency syndrome (AIDS) patients.

1.1.3. The hosts

Although medically important parasites, most *Leishmania* do not have humans as their most important host. Some species, like *L. arabica*, *L. turanica* and *L. gerbilli* in the OW and *L. aristidesi*, *L. enriettii*, *L. hertigi* and *L. deanei* in the NW have not been found in and are probably non-pathogenic to humans. Parasites that cause CL can easily be spread from person to person by the sandfly vector, since parasites are present in healthy skin, whilst human transmission of VL is only certain in India. Most *Leishmania* species have rodents and even dogs as reservoirs. Excluding humans, the main reservoirs for OW leishmaniasis by *L. major*, *L. infantum* and *L. aethiopica* belong to orders Rodentia, Carnivora and Hyracoidea, respectively. In the NW, orders Rodentia, Carnivora, Xenarthra and Didelphimorphia include reservoir species of *Leishmania*. Primates other than humans have been found infected with *L. shawi* in the NW and *L. major* in the OW (Binhazim *et al.*, 1987).

The dog is considered the main reservoir of VL caused by *L. infantum* and *L. chagasi*, although other canids have also been found infected (Dereure, 1999), like the red fox (Abranches *et al.*, 1984; Rioux *et al.*, 1968), the wolf (Dursonova *et al.*, 1965) and the jackal (Hervas *et al.*, 1996). *Leishmania infantum* has also been isolated from two Felidae (Ozon *et al.*, 1998) and four Rodentia species, including

two *Rattus spp* (Gradoni *et al.*, 1983; Ibrahim *et al.*, 1992). In the NW, *L. chagasi* was found to have two opossum species (*Didelphis spp.*) (Travi *et al.*, 1994) and the fox *Cerdocyon thous* (Lainson *et al.*, 1990; Lainson *et al.*, 1969; Mello *et al.*, 1989; Silveira *et al.*, 1982) as sylvatic reservoirs.

In the case of zoonotic VL (in the Mediterranean basin and the NW), humans are considered as dead-end hosts, although person to person spread is not completely ruled out. Notably, only human hosts are known for Indian VL, where parasites are easily found in peripheral blood and occurrence of PKDL can support the existence of a long term human reservoir. Although PKDL also occurs in the foci of Northeast Africa, evidence for an exclusive human reservoir is weaker, except perhaps in epidemic situations, but no reservoir has been positively identified. Acquisition of VL by people visiting game reserves, implicates the existence of a probably old, still unidentified, sylvatic reservoir. Thus, transmission of *L. donovani* in Africa may be of a mixed type, with human to human transmission, namely in epidemic situations, together with a well established sylvatic reservoir, which does not develop leishmaniasis and is infective to sandflies.

The large number of potential hosts of each *Leishmania* species complicates both evolutionary and epidemiological studies. Studies on co-evolution are hindered by the difficulty in identifying the most ancient reservoirs, whilst depletion or control of one reservoir may only switch the epicentre of the infection cycle to another host.

1.1.4. Immunity

In humans, self-healing cutaneous lesions are usually accompanied by positive delayed hypersensitivity skin reaction to leishmanin and low antibody production, whereas visceral leishmaniasis is usually accompanied by a surge in antibody production and the leishmanin skin test only becomes positive after cure (Ashford and Bates, 1999; Le Fichoux *et al.*, 1999). Diffuse cutaneous leishmaniasis, in which the immune response is hindered (Ashford and Bates, 1999), and the severe mucocutaneous leishmaniasis are much more difficult to treat (Lainson and Shaw, 1999; Ashford and Bates, 1999).

As *Leishmania* reach the skin of the mammalian host, as promastigotes, through the bite of a sandfly, they must face a series of defences by the immune system. *Leishmania*, however, can in part evade the immune system by parasitising macrophages. Promastigotes may differentiate extracellularly into amastigotes or may directly be phagocytosed by macrophages. Promastigotes efficiently activate complement and enter macrophages using complement receptor 3 (CR3) and other complement receptors (Mosser and Karp, 1999; Russell, 1995). Both

lipophosphoglycans and the protease glycoprotein 63 (GP63), abundant in *Leishmania* promastigote membranes, are known to bind CR3 directly and thus facilitate phagocytosis (Mosser and Karp, 1999).

1.1.4.1. In the macrophage

In the phagocytic vacuole, amastigotes are able not only to survive in the acidic environment, rich in lysosomal hydrolases, but transform it into a parasitophorous vacuole and avoid the macrophage antigen-presenting capabilities (Russell *et al.*, 1992). Although the membrane of the parasitophorous vacuole is rich in major histocompatibility complex type II (MHC II) molecules, *Leishmania* can reduce antigen presentation by limiting the release of potential antigens to their flagellar pocket (Russell *et al.*, 1992). However, both peptide and non-peptide antigens (such as lipoglycans) can be presented to human T cells, in the latter case by cluster of differentiation 1 (CD1) molecules.

In immature non-activated macrophages, amastigotes will proceed with division until they are released from the host cell to invade new macrophages. Mechanisms of immune evasion can vary according to the *Leishmania* species. Macrophages infected with *L. donovani* have lower levels of MHC II and co-stimulatory molecules (such as B7) and thus a lower capacity to present antigen to T cells (Russell, 1995), whilst *L. amazonensis* escapes the immune system by generation of certain epitopes (Russell, 1995), and there is evidence that *L. amazonensis* also internalizes and degrades MHC II molecules (De Souza Leão *et al.*, 1995). Furthermore, infected macrophages have a reduced ability to respond to activating cytokines, because of interference of the parasite with messenger cascades through modulation of protein kinase C (Russell, 1995).

Macrophages are activated by interferon γ (IFNγ) to produce inducible nitric oxide synthase (iNOs) which will synthesise reactive nitrogen radicals toxic to *Leishmania* (Stenger *et al.*, 1996). Nitrogen radicals such as nitric oxide (NO), however, seem to be able to control infection only in early stages of disease (Evans *et al.*, 1996). A functional Fas pathway is also necessary for resolution of disease (Stenger *et al.*, 1996). The cell surface Fas receptor mediates induction of apoptosis, via activation of caspases (cysteine proteases which cleave the carboxyl terminus of aspartate residues) and release of apoptogenic proteins (e.g., cytochrome c) from the intermembrane space of mitochondria to the cytosol, thus triggering degradative events (Wilson, 1998). During infection with *L. major,* macrophages up regulate surface expression of Fas in response to IFNγ, thus causing apoptosis (Stenger *et al.*,

1996). This reduction in the number of macrophages may limit the number of available host cells for parasite infection and also increase the ratio of CD4+ T helper 1 (Th1) cells, producing IFNγ, to infected macrophages. The efficiency of macrophage activation is thus increased (Stenger *et al.*, 1996). However, disruption of interleukin (IL) 2, a necessary potentiator of Fas-mediated cell death, or IL 2 receptor during infection may trigger autoimmune disease (Refaeli and Abbas, 1998). This mechanism might be involved in the development of the autoimmune type of mucosal disease due to *L. braziliensis*.

1.1.4.2. Innate resistance

The innate phase of host resistance is commanded by natural killer (NK) cells (Scharton Kersten and Sher, 1997), which can be activated upon *Leishmania* infection through both T-cell independent and dependent pathways (Scharton Kersten and Sher, 1997). Unlike T cells, NK cells respond very rapidly to stimuli and do not require priming (Stenger *et al.*, 1996). NK cells produce IFNγ and tumour necrosis factor α (TNFα) which can inhibit the growth of pathogens in initial stages of infection, thus allowing the host to develop an efficient adaptive immune response (Stenger *et al.*, 1996). Not only IFNγ has a role in activating macrophages, as NK cells influence the pathway of differentiation of CD4+ cells, through CD4, CD40, CD80 or CD86 and respective ligands.

1.1.4.3. T helper response

Correct activation of macrophages, leading to strong cell mediated immunity, is dependent on a number of Th1 type response cytokines, such as IFNγ and IL12 (Louis *et al.*, 1998) from CD4+ cells, and usually leads to cure (Gaafar *et al.*, 1995; Kurtzhals *et al.*, 1994). Active disease is dependent on a Th2 type response, with production of IL4 and IL10 (Cillari *et al.*, 1995; Kurtzhals *et al.*, 1994; Launois *et al.*, 1997; Reiner and Locksley, 1995); T-cells are depleted with consequent impaired lymphocyte and cytokine response to antigen stimulation. Th2 responses involve prominent humoral immunity with activation of antibody production (Kaye *et al.*, 1995; Kemp *et al.*, 1994). The outcome of infection has been shown to be dependent on the initial commitment of the immune system towards one of the response types (Louis *et al.*, 1998; Murphy, 1998; Reiner and Locksley, 1995) in which case relative concentrations of Th1 and Th2 type cytokines are more important than absolute quantities of cytokines. This model is mainly based on murine studies, but similar polarisation of the immune response seems to control the response to *Leishmania*

infection in humans (Kemp *et al.*, 1994; Kurtzhals *et al.*, 1994).

Development of a Th1 type response is dependent on a number of factors, inlcuding time after infection. IL12 is critical but only in the initial stages of infection and may thus be useful as a vaccine adjuvant rather than as therapy (Scharton Kersten and Sher, 1997). IFNγ can regulate the ability of T cells to respond to IL12, but is not enough for Th1 development (Murphy, 1998). However, IFNγ produced by specific MHC II restricted CD4+ T cells is important in regulating macrophage activation (Louis *et al.*, 1998) by activating iNOs and up-regulating surface Fas. *Leishmania* can specifically inhibit IL12 production by infected macrophages *in vitro* and thus lead to the establishment of chronic patent infections (Carrera *et al.*, 1996) and may cause disseminated VL in humans by engaging Fcγ receptors on macrophages, which stimulate IL10 production (Karp *et al.*, 1993).

IL4 is the main cytokine implicated in Th2 development, with the simultaneous loss of IL2 responsiveness (Murphy, 1998). Th2 cytokines deactivate macrophages and hamper the effect of IFNγ; furthermore, tumour growth factor β (TGFβ), IL10 and IL13 can interfere with induction of iNOs (Stenger *et al.*, 1996).

Resolution of infection may not mean freedom from infection. Inhibition of the nitric-oxide-dependent pathway can result in expansion of parasites that had been either quiescent or replicating slowly in resolved lesions of *L. major* in mice (Louis *et al.*, 1998). Furthermore, maintenance of delayed type hypersensitivity, or positivity to the Montenegro test, and the many cases of PKDL after resolution of leishmaniasis suggest non-sterile immunity.

1.1.4.3. External factors regulating immunity

Host resistance to infection appears to be controlled by several genes and may involve several mechanisms (Stenger *et al.*, 1996), as described above. Despite the importance of the host response, external factors, such as species of *Leishmania*, dosage, number of inoculations and vector saliva, can influence the outcome of infection.

Several studies have revealed that sandfly saliva may affect host susceptibility to *Leishmania* infection. Saliva components were shown to enhance lesion development by *L. major* (Belkaid *et al.*, 1998; Hall and Titus, 1995; Mbow *et al.*, 1998) and *L. braziliensis* (Donnelly *et al.*, 1998; Samuelson *et al.*, 1991), and to affect the outcome of *L. infantum* infection (Warburg *et al.*, 1994). The saliva of sandflies is highly immunogenic (Ghosh and Mukhopadhyay, 1998) and previous exposure of the host to the saliva may reduce infectivity of *Leishmania* (Belkaid *et al.*, 1998), although

saliva was shown to decrease killing of intracellular parasites (Hall and Titus, 1995) and to exarcebate disease by *L. major* by decrease of Th1 cytokines and associated factors and enhancement of Th2 cytokine IL4 (Mbow *et al.*, 1998), the latter even in the absence of infection. The levels of a salivary component, maxadilan (Lerner and Shoemaker, 1992), were shown to correlate directly with erythemas caused by feeding sandflies and inversely with *Leishmania* proliferation in the skin (Warburg *et al.*, 1994). This effect could explain proliferation in the cutaneous lesions produced by *L. chagasi* infection in Costa Rica, where the population of *L. longipalpis* has a lower level of Maxadilan than Brazilian or Colombian sandflies (Warburg *et al.*, 1994).

1.1.4.5. Vaccination

There is some evidence that protection against leishmaniasis is possible; induced infection with *L. major* is still practised to protect against undesirable face lesions (Modabber, 1990) and there are a large number of asymptomatic leishmanin positive individuals in endemic areas (Marty *et al.*, 1992; Marty *et al.*, 1994; Nandy *et al.*, 1987; Shiddo *et al.*, 1995; Zijlstra *et al.*, 1994). Long-term protection seems to be associated with CD8+ producing IFNγ (Gurunathan *et al.*, 1997; Pinelli *et al.*, 1994), a component overlooked in most vaccine development trials. (For types of vaccine and vaccination trials see 1.1.6. Control of VL)

1.1.5. Diagnosis

Clinical presentation of visceral leishmaniasis, especially at the onset, can be, and has been, confused with a number of other diseases, including malaria, liver diseases and immune disturbances (DeBeer *et al.*, 1991). Even the more conspicuous cutaneous forms of leishmaniasis can be mistaken for leprosy or other cutaneous diseases (Blum *et al.*, 1994; Chakrabarti *et al.*, 1997; Dhar *et al.*, 1995; Ramesh *et al.*, 1994). It is thus necessary to use specific diagnostic methods. The gold standard has been and still is microscopical identification of parasites from biopsies of infected tissue (Palma *et al.,* 1991). Specificity is 100%, but sensitivity is low and dependent on preparation and examination of slides. Diagnosis can also be made by isolation of parasites in culture or by sub-passage in a rodent. These methods increase sensitivity, but are demanding in terms of cost, time and manipulation. Serological methods allow indirect identification of infection (or exposure to infection) by measuring the amount of specific antibodies against *Leishmania* antigens. The most used serological techniques are direct agglutination tests (DAT) (el Harith *et al.*, 1989), immunofluorescence (IFAT) (Kien Truong *et al.*,

1969; Lupascu *et al.*, 1970) and enzyme-linked immunosorbent assay (ELISA) (Hommel *et al.*, 1978), although others are also available, such as Western blot (dos Santos *et al.*, 1987; Jaffe *et al.*, 1984), dot-ELISA (Pappas *et al.*, 1983), latex agglutination test (Mayrink *et al.*, 1972), immunodiffusion (Bray and Lainson, 1966; Rodriguez Cuartero and Nunez Carril, 1974), complement fixation (Chavez and Witremundo Torrealba, 1965) and counterimmunoelectrophoresis (Desowitz *et al.*, 1975). Serological methods are particularly useful in the diagnosis of visceral leishmaniasis, since the active phase of disease is characterised by increased humoral response which decreases with cure, but less for cutaneous forms of leishmaniasis with have a stronger cellular response. Specificity and sensitivity are not ideal, however, because of low antibody titres and cross-reactivities, and diagnosis is usually improved by combining two or three different methods, for example IFAT, ELISA and DAT (Cabral *et al.*, 1998; Cuba *et al.*, 1996; Kar, 1995; Millesimo *et al.*, 1996; Semião Santos *et al.*, 1995).

DNA based methods for detection of parasites, which have been developed to be 100% specific and can be very sensitive (Degrave *et al.*, 1994; Weiss, 1995; Wilson, 1995), include detection by hybridisation, with DNA probes, of whole or restricted DNA on several types of blots, such as dot blots and Southern blots. The most promising are the polymerase chain reaction (PCR) and its derivatives (Degrave *et al.*, 1994; Weiss, 1995; Wilson, 1995). In practice, DNA based diagnosis is not free from problems: lack of effective sensitivity, false positives, inhibition by host DNA or contaminants, sensitivity to contamination by amplicons, etc. PCR technology is recent and it has been difficult and expensive to apply in field conditions or in less developed countries, where leishmaniasis is a bigger problem. In conclusion, despite some advantages of molecular biological diagnostic methods, serological methods are still the best choice in most countries, although DNA based methods are becoming cheaper and easier to use.

Optimal diagnosis of leishmaniasis should involve identification of the parasite species or even phenotype, because different parasites can cause identical disease forms and, conversely, the same species of parasites can cause different disease forms. The best treatment may vary according to parasite types (Berman, 1997; Chance, 1995; Davidson, 1998). In epidemiological surveys, it is not only essential to know how many people are infected but also with which parasite, in order to understand life cycles and develop better control measures. At present it is in this field that molecular methods are the most useful, since serological methods are inadequate and current methods of allozyme electrophoresis are time consuming and require large numbers of parasites. DNA based typing should be faster, more reliable

and provide characterisation of the parasites at different taxonomic levels, according to the needs of the epidemiologist or clinician.

1.1.6. Control of visceral leishmaniasis

A parasitic disease with complicated epidemiology such is leishmaniasis, is very difficult to eradicate (Wilson and Streit, 1996). Although apparent eliminations have been achieved, often through vector or reservoir control, re-introductions or resurgences are very common (see 1.1.2.). The best examples of successful eradications may have been in some areas of China (Bao *et al.*, 1994) and Iraq, but knowing the present situation is very difficult, especially in Iraq. VL remains in some areas of China (Katakura *et al.*, 1998) and *L. infantum* has been found in sandflies in areas where only CL had been reported (Guan *et al.*, 1994). Some *L. infantum* strains are known to cause CL, but *L. infantum* may also cause CL or inapparent infections in populations with better health conditions.

Vector control using insecticides has been effective to some extent but it also damages the environment. Given the zoophylic and burrow breeding habits of sandflies, large areas should be sprayed, not only humid or habitational areas. Large amounts of insecticide would thus be required, raising cost and contamination of human and natural environments. Resurgence of sandflies has occurred after major depletions in number, often after reported they were thought to have disappeared, following insecticide anti-malarial campaigns, probably because they are more resistant than *Anopheles* to dry environments, which are often spared spraying, and because PKDL patients may act as long term reservoirs (Thakur and Kumar, 1992; Wilson and Streit, 1996). In most cases resurgences were made worse when populations became immunologically naive, like in Bangladesh (Elias *et al.*, 1989) and India (Thakur and Kumar, 1992). There is also some evidence that sandflies may develop insecticide resistance (Bansal *et al.*, 1996; Kaul *et al.*, 1994; Mukhopadhyay *et al.*, 1996), thus requiring increasing amounts of or alternative insecticides.

Reservoir control has been done by culling or treatment of sick or infected animals, mainly dogs. Treatment of human patients is mostly effective, as a control measure, against antroponotic VL (as in India) together with the use of bednets to prevent spread to the vector. For zoonotic VL, treatment of dogs is not recommended since it is largely ineffective and very expensive, and it is not an option in impoverished populations. Although identification and killing of infected dogs has been mostly applied to control VL with success in some areas, such as China and Iraq, in most countries it is very difficult to achieve due to the logistics of catching stray dogs, identifying all infected dogs and convincing owners to sacrifice their dogs.

Culling infected dogs may not be effective because at least some dogs may be infective to the sandfly before the onset of clinical signs and a significant serological response (Molina *et al.*, 1994). Other reservoirs may be wild or semi-domestic animals. It is thus very difficult to identify all infective reservoir animals. Furthermore, in some areas the human reservoir may become more important as cases of HIV / *Leishmania* co-infection increase and create a human restricted cycle of leishmaniasis.

Vaccination is the most sought for form of control, for it should be more cost effective and more efficient. Complete eradication may never be achieved, but maintenance of very low infection levels in resistant populations would be highly desirable and a realistic objective. Historically, live vaccination with *L. major* has been used with success to prevent disfiguring lesions, however, it is both impractical and dangerous, and vaccination with *L. major* does not cross protect against other *Leishmania* species. Indeed, there is very limited cross-protection between *Leishmania* species. Trial vaccination with killed whole parasites of *L. infantum* increased susceptibility of dogs in the field (Dunan *et al.*, 1989). Second generation vaccines involve identification and isolation of antigens. Although many *L. major* antigens have been successful in the rodent model against cutaneous leishmaniasis, effective protection against *L. infantum* or *L. donovani* in the dog has proved elusive. Third generation vaccines include recombinant DNA, either in other organisms, such as BCG, *Salmonella*, etc. or as naked DNA in plasmids. Furthermore, vaccination against VL in the Mediterranean basin and the Americas would be most effective on the dogs, the reservoir. Currently, in Brazil there is a promising candidate whole vaccine which is being tested in phase III trials (Mayrink *et al.*, 1996).

Good vaccination strategies against leishmaniasis are difficult to develop. The mouse is a model preferred to the dog for technical and ethical reasons, but the two models are not fully compatible. Immunology of leishmaniasis is poorly understood and dependent on the parasite species and the host and vaccine candidates can only be identified by trial and error. A suitable antigen should elicit both cell mediated immunity and memory. In this respect, DNA immunisation could be the best option. An alternative could be to vaccinate against vector saliva (Ghosh and Mukhopadhyay, 1998), as sandflies fed with immunized serum or from previously exposed animals have a higher death rate and development of *Leishmania* in the vector is affected.

The main risk factor for human leishmaniasis seems to be malnutrition and immune deficiencies. Hunger, as a consequence of poverty, may never be eradicated and malnutrition, either through poverty or fast food culture in rich countries, is

frequent. The number of immune deficient patients is increasing due to AIDS and transplant patients and more and more immunologically naive people migrate into leishmaniasis areas.

It is very likely that neither eradication, nor even the perfect strategy for control of leishmaniasis will ever be achieved. Not only, in many cases, is the source of infection present in the wild, like in the Amazonian forest, but also in other cases the reservoir is becoming more and more domestic - the dog, or even the rat. Many reservoirs are not exclusive to one species of parasite and many parasites adapt to new hosts (as the dog) as the sandfly vector changes feeding habits. With this in view, the best strategy for control of leishmaniasis will probably be a combination of a number of methods, locally adapted to the parasites, vectors and reservoirs and economic and ecological characteristics of each focus, but in a global control strategy.

A good summary of the present situation is presented by Wilson and Streit (1996): "Visceral leishmaniasis presents a serious problem in endemic regions that is difficult to treat or prevent. Several epidemiological problems make the disease particularly troublesome to manage. These include the facts that classic visceral leishmaniasis is fatal if untreated and there is not reliable access to medical care in many endemic regions. When available, treatment has associated toxicity and requires the use of intravenous medications with careful monitoring for toxicity, which are complex to administer in underdeveloped nations. There is an increasing incidence of the disease in HIV-infected individuals in southern Europe, in part because of the fact that eradication of the organism from infected persons using currently available drugs appears to be difficult if not impossible. Furthermore, chronic cutaneous forms of the disease allow humans and animals to maintain the organism long-term in a bodily site that is easily accessible to the sandfly vector. More effective and less toxic treatment modalities as well as a protective vaccine are badly needed to manage this disease."

1.2. The parasite

1.2.1. Characteristics and morphology

Leishmania are eukaryotic single celled organisms, with a modified mitochondrion of which the genome (see 1.2.5.2.) is organised in a disk shaped dense structure called the kinetoplast which is located near the base of the flagellum. Parasites are found in two basic forms: the promastigote which is characterised morphologically by an elongated body of 5 to 20 μm in length and 1 to 4 μm width, and an anterior free flagellum of up to 20 μm; and the round or oval amastigote, of

4μm in length and 2 μm wide on average, with a very short, almost invisible, flagellum enclosed in the flagellar pocket (Antoine *et al.*, 1999). A third form, the paramastigote (Lang *et al.*, 1991) with only 5 to 6 μm of length and a very short flagellum, has been described in the vector but the role in the life cycle is not known (Antoine *et al.*, 1999).

Leishmania are kinetoplastids, considered to be an ancient group descendent from primitive mitochondriate eukaryotes (Vickerman, 1994), although possessing most of their characteristic structures (Antoine *et al.*, 1999). The flagellum arises from the eukaryotic basal body in the base of an anterior invagination of the cell membrane - the flagellar pocket - and has an eukaryotic axoneme (9+2) structure, although a paraxial rod of paracrystalline structure is also present which accompanies the axoneme along its entire length (Vickerman, 1974). The basal body has a companion second basal body, in a centriolar construction (Vickerman, 1974). The flagellum of trypanosomatids has the unique ability to form junctional complexes of the desmosome or hemidesmosome type with the outer membrane of its own cell or with other flagellates, or the substrate, notably the insect gut lining (Antoine *et al.*, 1999; Vickerman, 1974). The flagellar pocket is the only membrane portion capable of fusion and thus of endo or exocytosis (Antoine *et al.*, 1999). The kinetoplast as seen by electron microscopy, is a network of circular DNA molecules (see 1.2.4.) which is localised in the interior of the single mitochondrion near the base of the flagellar pocket (Vickerman, 1974). The mitochondrion has a typical double membrane structure with cristae on the inner membrane and is often found forming a complex labyrinthic net with the endoplasmic reticulum (Vickerman, 1974). The Golgi apparatus is readily recognisable, with its multivesicular bodies at its distal ends. A pulsatile or contractile vacuole empties from this region into the flagellar pocket (Vickerman, 1974). Other eukaryotic organelles such as lysosomes are present (Antoine *et al.*, 1999), but also a typical kinetoplastid organelle, the glycosome, which is related to peroxisomes and glyoxysomes in high eukaryotes, with important metabolic functions (Antoine *et al.*, 1999). Another peculiar organelle type is the acidocalcisome (Vercesi *et al.*, 1994), also present in apicomplexan parasites, which is very rich in cations (phosphates, Ca^{2+}, Mg^{2+}, Na^{+}) (Docampo and Moreno, 1999; LeFurgey *et al.*, 1990) and membrane pumps (Docampo *et al.*, 1995; Vercesi *et al.*, 1994). Its function is unknown but this organelle may regulate the *Leishmania* internal pH (Docampo and Moreno, 1999), store Ca^{2+} for signalling infective stages (Lu *et al.*, 1997) or store energy (Docampo and Moreno, 1999; Scott *et al.*, 1997).

The *Leishmania* promastigote membrane is very rich in lipophosphoglycan and a glycosyl-phosphatidylinositol (GPI) linked metalloenzyme, glycoprotein 63

(GP63), also known as major surface protease (msp), promastigote surface protease or leishmanolysin (see 6.1.2.), both of which are very reduced or disappear in the amastigote phase (Antoine *et al.*, 1999). These molecules play important roles in the life cycle of the parasite. Glycoinositol-phospholipids are present in both stages of the parasite and may have a structural role or protect from the low pH in the parasitophorous vacuole or hydrolases in either stage. Other membrane proteins include ATPases (Meade *et al.*, 1989) and those in the polymorphic families of M-2/PSA-2 (Murray and Spithill, 1991; Murray *et al.*, 1989) and protein B (Flinn *et al.*, 1994; Pimenta *et al.*, 1994).

Cell division in *Leishmania* begins with replication of the flagellum (Antoine *et al.*, 1999; Vickerman, 1974). The kinetoplast divides simultaneously with the mitochondrion and before the nucleus. *Leishmania* do not have classic mitosis because they lack visible chromosomal condensation, although they have stable chromosomes. The nucleus is enveloped in a microtubule net and elongates to divide in two (Vickerman, 1974). It is not clear whether the chromosomes are attached to microtubules or to the nuclear envelope (as in dinoflagellates). Meiosis is unknown as are any sexual stages, although fusion of parasites has been observed (Lanotte and Rioux, 1990) and hybrids have been identified (Belli *et al.*, 1994; Delgado *et al.*, 1997; Dujardin *et al.*, 1995a; Kelly *et al.*, 1991).

1.2.2. Life cycle(s)

Leishmania are intracellular parasites of macrophages in a wide range of mammalian hosts which are transmitted through the bite of insects of the genus *Phlebotomus* or genus *Lutzomyia*. Highly motile promastigotes accumulate in the foregut of the sandfly and are expelled during the insect's blood meal (Walters, 1993). In the mammal, promastigotes are phagocytosed by macrophages (Ashford and Bates, 1999) or may differentiate into amastigotes (el Azzouni *et al.*, 1998), which are then phagocytosed. Low pH or temperature increase (34-36°C) were shown to trigger amastigote-like differentiation (Bates, 1994; Pan *et al.*, 1993). The phagocytic vacuole becomes a parasitophorous vacuole (Antoine *et al.*, 1998), where amastigotes divide and eventually are released to infect other macrophages. Infected macrophages can remain in the skin where they are eventually eliminated (most CL) or amastigotes can be carried to internal organs (VL) or to the mucosae (MCL) where they can remain until treatment or death of the host.

Sandflies may ingest infected macrophages with a blood meal. The amastigotes are released and differentiate into procyclic promastigotes which are

short, actively dividing and have low motility, and attach to the gut epithelial cells (Killick Kendrick and Rioux, 1991). The insect often takes a sugar meal between two consecutive blood meals which enables differentiation (metacyclogenesis) into infective metacyclic promastigotes which are long, non-dividing, highly motile and migrate to the mouth parts to start the cycle again.

The different morphological forms can not only be differentiated by shape and behaviour, but also by surface antigen composition (Moody, 1993). Promastigotes present a thick glycocalyx rich in GPI-anchored glycoconjugates, which can be divided into two classes: glycolipids and glycoproteins. The main glycolipid is a lipophosphoglycan, but there are also glycosylinositol phospholipids (see 1.2.1.).

1.2.3. Evolutionary history of *Leishmania*

1.2.3.1. The kinetoplastids

Kinetoplastids are peculiar eukaryotes with characteristic features, such as a kinetoplast, RNA editing, a glycosome, a nucleus with polycistronic transcripts. The surface membrane, the cytoskeleton, the flagellar pocket are also characteristic in kinetoplastids. They differ from cilliates in that they combine functional diversity with morphological conservation (Vickerman, 1994).

It is accepted that kinetoplastids are related to free living euglenoids, but opinions are divided as to the origin of parasitism (Vickerman, 1994). The group includes monogenetic as well as digenetic parasites and hosts include fishes, amphibians, mammals, reptiles and plants. Vectors are usually insects but leeches transmit fish trypanosomes. It is far from clear whether kinetoplastids evolved with insects and were accidentally transmitted to vertebrates, which became hosts, or first evolved with vertebrates and secondarily acquired transmission through insects. Evidence supports both hypotheses but knowledge of the biology of species non-infective to man is poor. It is also highly likely that most kinetoplastid species have not been found yet, let alone studied, especially free living species and in marine animals and reptiles. In most phylogenetic trees of Kinetoplastida, parasitism either has arisen independently in different lineages (Vickerman, 1994), or has been lost in others.

Kinetoplastids are metabolically very versatile and can use both amino acids and sugars, thus facilitating adaption to new hosts. Different life cycles and different hosts may mean that reconstruction of phylogenies based upon transmission data can become very difficult with different evolution rates and (or) absence of co-evolution.

Vectors and modes of transmission of the different groups are extremely divergent and, although kinetoplastids may infect insects, development of vectorial

capacity was probably acquired independently. *Trypanosoma brucei* and *Leishmania* are both transmitted by Diptera but by sub-orders Brachycera and Nematocera, respectively, and *T. cruzi* is transmitted by a different Order of insects (Hemiptera).

The most important human parasites within the Kinetoplastida, other than *Leishmania*, are the trypanosomes, which seem to have had a common ancestor and may have differentiated into *T. brucei* and *T. cruzi* clades during the split between South America and Africa (Alvarez *et al.*, 1996; Stevens and Gibson, 1999; Wright *et al.*, 1999). However, a trypanosome 'aquatic clade' has also been identified, somewhat closer to the *T. cruzi* clade, in terms of genetic distance, and the evolutionary relationships with the other two clades are obscure. In the most comprehensive tree of the trypanosomatids obtained so far, a *Leishmania / Crithidia / Endotrypanum* clade was evidenced (Stevens and Gibson, 1999). Definition of relationships within this clade has been difficult and it is possible that the group suffered a rapid radiation, making true phylogenies difficult to detect.

1.2.3.2. *Leishmania*

The evolutionary history of *Leishmania* seems to be connected with that of strains classified as *Endotrypanum*, described as an intra-erythrocytic parasite (Shaw and Bird, 1969). Some *Leishmania* species (*L. colombiensis*, *L. deanei, L. equatoriensis*, *L. herreri* and *L. hertigi*) were found to be part of an *Endotrypanum* clade (Cupolillo *et al.*, 2000) and are not known to infect man, except *L. colombiensis*. A redefinition of the *Leishmania* genus was proposed recently, with division into an Euleishmania Section with true *L.* (*Viannia*) and *L.* (*Leishmania*) sub-genera and a Paraleishmania Section with *L. colombiensis*, *L. deanei, L. equatoriensis*, *L. herreri*, *L. hertigi* and current laboratory strains under the genus *Endotrypanum* (Cupolillo *et al.*, 2000). The *Endotrypanum* clade, which has only been found in the NW, was very close to the Euleishmania Section and the sub-genus *Viannia*, also only found in the NW, is usually close to the root of the *Leishmania* clade. Therefore, a NW origin for the *Leishmania* genus was proposed (Noyes, 1998b), although other alternatives, such as a Paleartic (Kerr, 2000) and an African origin (Momen and Cupolillo, 2000) have also been defended.

OW *Leishmania* seem to have less genetic diversity, on the whole, than NW *Leishmania*, which suggests that they may be of more recent origin. NW *Leishmania* could have reached the OW through Beringia during the Miocene, perhaps transported by infected critecid rodents which underwent a major radiation and migration at that time (Noyes, 1998b). The most likely ancestor of OW species is the

L. mexicana complex, which causes CL similar to that by *L. major. Leishmania major* like strains have been identified in the Americas, although their origin is dubious. Independently of their ultimate origin, most OW *Leishmania* species occur in Northeast Africa (Sudan, Kenya, Ethiopia) or the Near / Middle East, where they may have originated and from where they may have spread. These are regions where humans and later agriculture may have first developed. Some Chinese *L. donovani* and *L. infantum* are genetically divergent from those of other regions which could be a local phenomenon enhanced by genetic isolation and new ecological niches or those strains could be remnants from early OW colonists.

Some of the most intringuing *Leishmania* are the blood parasites of reptiles which have been found exclusively in the OW. These parasites had been regarded as subgenus *L.* (*Sauroleishmania*) but were classified more recently in a different genus *Sauroleishmania* Ranque, 1973 (Killick-Kendrick *et al.*, 1986). In contrast, phylogenetic analyses of *Leishmania* usually place the *Sauroleishmania* clade well within the genus *Leishmania* (Croan *et al.*, 1997; Noyes *et al.*, 1997), which provides an additional difficulty for the search of the origin of *Leishmania*. If *Leishmania* originated in the NW, then, how to explain the frequent branching of *Sauroleishmania* before the sub-genus *Leishmania* (including *L. mexicana*)? On the other hand, if *Leishmania* originated in the OW, how to explain the greater genetic diversity in the NW and the relationship with the NW specific *Endotrypanum* clade? Noyes (1998b) suggested that burrow sharing between *Leishmania* infected rodents and reptiles may have facilitated cross transmission by the sandfly co-inhabitants if the insects were not specific feeders. Parasites might even have been initially transmitted through direct blood contact, for example in fights, or by feeding on infected sandflies. Those parasites which successfully had infected lizards would undergo faster evolution than the mammalian relatives through different selection pressures in a very different host. The faster evolution would explain the early branching patterns obtained in some trees of the group by a long branch attraction effect, if they are relatively recent. There is some evidence that the genetic divergence rates of *Sauroleishmania* are higher than among other *Leishmania* (Croan *et al.*, 1997). Those parasites are transmitted by a different genus of sandflies, but the phylogeny of the vector is only poorly understood. It has been found that mammalian *Leishmania* can be infective to lizards (Noyes *et al.*, 1997), thus supporting a close relationship between *Leishmania* and *Sauroleishmania*. Furthermore, lizard *Leishmania* were found to be capable of causing transient infections of mammals (and humans) (Adler, 1964, Belova, 1971, Manson-Bahr, 1961 cited in Noyes *et al.*,

1997). If *Sauroleishmania* are closer to OW *Leishmania*, then the genus may be paraphyletic and *Sauroleishmania* need to be considered as a sub-genus of *Leishmania* again.

Phylogenetic reconstruction of *Leishmania* is hampered by lack of knowledge of their ecology and also by their diverse range of hosts. Higher molecular diversity found in natural populations of a given *Leishmania* species was found to be related with the higher number of sandfly vector(s) and / or animal reservoir(s) involved in the transmission cycle of the parasites in a possible, although not strict, co-evolution phenomenon (Cupolillo *et al.*, 1998). Generation time is very difficult to determine as the life cycles in the natural reservoirs (when known) remain to be fully elucidated. All known *Leishmania* are transmitted by sandflies as far as it is known with infection restricted to the digestive system, although excretion with urine has been demonstrated (Sadlova and Volf, 1999). The importance of this finding is difficult to balance at the momment, although it suggests a possible alternative means of transmission (including vector to vector) without a blood feeding cycle.

1.2.4. Current taxonomy of *Leishmania*

The classification of Hausmann and Hulsmann (1996) includes *Leishmania* in the Empire Eukaryota, Kingdom Mastigota, Subkingdom Dimastigota, Superphylum Metakaryota, Phylum Euglenozoa, Subphylum Kinetoplastida, Class Trypanosomatidae, Genus *Leishmania* Ross, 1903. However, most researchers still use the classification by Levine *et al.* (1980), particularly below and at the level of Order, which includes *Leishmania* in the Kingdom Protista Haeckel, 1866, Sub-kingdom Protozoa (Goldfuss, 1817) Siebold, 1848, Phylum Sarcomastigophora Honigberg & Balamuth, 1963, Sub-phylum Mastigophora Diesing, 1866, Class Zoomastigophorae Calkins, 1909, Order Kinetoplastida (Honiberg, 1963) Vickerman, 1976, Sub-order Trypanosomatina Kent, 1880, Family Trypanosomatidae (Doflein, 1901) Grobben, 1905, Genus *Leishmania* Ross, 1903.

The classification of protozoa is not consensual and variations in the relative categories of each group and the higher groups can vary among authors (Corliss, 1998). The protozoa are a very difficult group to work with because of the lack of evident morphological structures and varied life styles. The long time elapsed since major divisions have occurred makes it difficult to use general molecular clocks because of the lack of homologies in many cases and the accumulation of mutations beyond an optimal resolution level.

If higher levels of classification are difficult to establish, similarly, the species

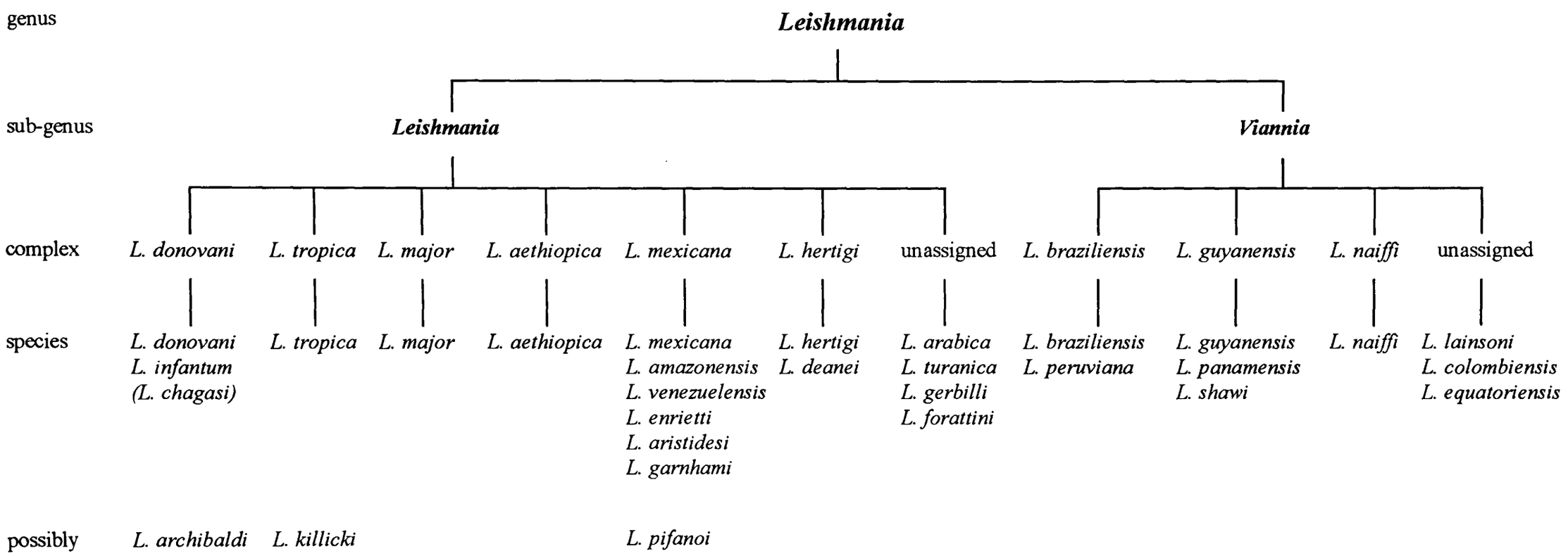


Figure 6 - Summary classification of the *Leishmania* (Rioux *et al.*, 1990; Shaw, 1994; WHO, 1990). Many assignments are controversial and may vary according to different authors.

level is not any easier. The biological concept of species is impossible to apply to asexual organisms, since phylogenetic trees obtained from molecular data are probably often a collection of clonal lines. The clonal pattern can often be complicated by occasional hybridisation and reticulate evolutionary trees are very difficult to interpret if forced into dichotomous trees.

There are a number of major branches within *Leishmania*, which have been classified as sub-genus, complexes of species, species, sub-species. The nomenclature is far from stable. Although these branches are easy to identify with a small number of strains per group, larger samples make trees much more difficult to interpret and it is often apparent how close the most ancient strains are to each other, even if they are from different 'species'.

Other genera in the Trypanosomatidae are *Crithidia*, *Leptomonas*, *Herpetomonas*, *Blastocrithidia*, *Trypanosoma*, *Phytomonas* and *Endotrypanum*. The genus *Sauroleishmania* has been established for *Leishmania* like parasites found in the blood of reptiles. Those parasites were initially classified as *Leishmania* sub-genus *Sauroleishmania* but were judged as sufficiently different to belong to a different genus. Molecular classifications, however, have placed some of the *Sauroleishmania* species within the *Leishmania* group. A better understanding of the reptile *Leishmania* is not possible with the limited number of isolated strains.

The genus *Leishmania* has been divided into two sub-genera, according to the development in the insect gut (Lainson and Shaw, 1987); the *Viannia* include a phase of division in the hindgut of the sandfly and the *Leishmania* are limited to the foregut and midgut. This rather arbitrary division has, however, been confirmed by molecular data. The *Viannia* subgenus is limited to the NW but the *Leishmania* sub-genus is spread through the whole distribution of leishmaniasis.

In an attempt to rationalise the complex classification of *Leishmania* species, they were assigned into complexes of species or strains which reflect monophyletic groups with similar epidemiology and clinical presentation (Fig. 6). A number of species remain unassigned.

Traditionally, *Leishmania* have been differentiated by allozyme (isoenzyme) analysis and different phenotypes sharing certain profiles are named zymodemes. The degree of genetic diversity within each species varies enormously; some are extremely polymorphic and others are very conserved. This is a consequence of the traditional designation of species according to geographical distribution, clinical presentation, epidemiological characteristics. The limited number of enzymes chosen for analysis and the system used do not allow full exploration of the extent of diversity and enzyme based classifications are necessarily biased (see chapter 3).

1.2.5. Genetics of *Leishmania*

Leishmania have two genomic pools: the nucleus and the kinetoplast, which corresponds loosely with the classical mitochondrial genome of higher eukaryotes.

Viruses have been found and characterized in *Leishmania*, mostly in *L. braziliensis* and *L. guyanensis* and also one in *L. major* 5-ASKH (Widmer and Dooley, 1995), but virus-like particles have been reported in *L. hertigi* (Molyneux, 1974). The *Leishmania* viruses have ~ 5280 base pairs (bp) double stranded RNA within a spherical capsid and seem to be close relatives of the yeast L-A/ScV virus. Phylogenies of *Leishmania* infected strains and respective viruses suggested a long association time between the two (Widmer and Dooley, 1995), but this hypothesis is difficult to reconcile with present hypotheses for *Leishmania* evolution.

1.2.5.1. Nuclear genome

The nuclear genome is arranged as linear chromosomes that do not undergo condensation. For this reason, study of the *Leishmania* chromosome organisation was not possible until pulse-field gradient gel electrophoresis (PFGE) was developed. This electrophoretic technique was applied initially to the smaller chromosomes of yeast and revolutionised kinetoplastid genetics. The direction of the electrical field during electrophoresis is changed at regular times thus allowing the separation of the full length chromosome-sized DNA.

The *Leishmania* nuclear genome has a GC content of about 60% (Alvarez *et al.*, 1994) and contains 36 chromosomes (Wincker *et al.*, 1996), of between 200 and 4000 kilo base pairs (kb) in length (Lighthall and Giannini, 1992). *Leishmania* seem to be mainly diploid (Bastien *et al.*, 1992; Cruz *et al.*, 1991), although some strains might be aneuploid because extensive size differences could be identified between some homologous chromosomes. There may be extensive size differences between homologous chromosome pairs of different species and karyotypes may be used to distinguish between some species. Because of the size variability of *Leishmania* chromosomes, they have to be identified by probe hybridisation and linkage studies (Wincker *et al.*, 1997). The assignment of specific markers to chromosomes facilitates mapping of genes as part of the genome sequencing project. Telomeres have been identified, and are the origin of much of the chromosome size variability (Wincker *et al.*, 1996), but centromeres have not been found.

Genes are often present in tandem arrays, or as multiple copies; gene amplification was once thought of as being responsible for chromosome size polymorphisms. Genes, so far, lack introns and multicistronic transcription starts at a

single 5' promoter (Graham, 1995). Mature RNAs are generated by co-transcriptional capping, through trans-splicing (Graham, 1995; Nilsen, 1994). Small capped non-polyadenylated RNAs, consist of a 5' exon (Nilsen, 1994) (the spliced leader or mini-exon) and a 3' splice donor site, are produced and associate with each forming messenger RNA (mRNA). The spliced leader is linked to the 5' end of the transcribed region of the mRNA (Graham, 1995; Nilsen, 1994). This process caps the mature mRNA, and polyadenylation is thought to be functionally coupled to trans-splicing (Nilsen, 1994) during multicistronic transcript processing .

1.2.5.2. Kinetoplast genome

The kinetoplast genome is composed by a complex network of circular DNA. In the Kinetoplastida there are approximately 20-50 maxicircles of 20 to 40kb and 5 x 10^3 - 10^4 minicircles (Simpson, 1987) of 0.5 to 2.9kb, making up to 90% of the mass of the kinetoplast (Shapiro and Englund, 1995). Maxicircles are the functional equivalent of mitochondrial DNA of other eukaryotes (Shapiro and Englund, 1995). They encode mitochondrial ribosomal RNA (rRNA) and proteins, most of them metabolic enzymes, and have a variable region within the replication unit. Minicircles have extremely variable sequences except for a region of 100 to 180 bp which contains the origin of replication, the rest of the sequence encodes guide RNAs (Shapiro and Englund, 1995). The organisation of the kinetoplast is complex with each minicircle linked to three neighbouring minicircles in a tight disk shaped network. Replication is done with the aid of two complexes of replication proteins which are located at opposite edges of the disk. Minicircles are freed from the centre of the disk and migrate to the edge where they are replicated and become attached to the disk once again. Maxicircles are encased in the disk and replicate attached to it. Otherwise, replication of both type of circles seems to be dependent on the same initiation sequence and probably involves the same enzyme mechanism. Replication of the kinetoplast network involves replication of the DNA to twice the initial number of molecules and remodelling of the minicircle distribution within the structure, followed by scission of the network and a second remodelling to create two separate and fully organised kinetoplasts. There seems to be a strong selective pressure to ensure that all classes of minicircles are transmitted to both daughter cells in most kinetoplastids. Minicircles have an important role in maxicircle gene expression and little variation is observed in restriction patterns after several culture subpassages of *Crithidia fasciculata* and *Sauroleishmania tarentolae*. Kinetoplast DNA uniqueness does not stop at the organisational level. Maxicircle transcripts undergo extensive

RNA editing, with addition or deletion of uridine residues at sites specified by guide RNAs synthesised from minicircles.

1.2.5.3. Genes

Leishmania have, as kinetoplastids, some characteristic genes which are not present in other eukaryotes, like trypanothione reductase (Keithly, 1989; Shames *et al.*, 1986) and mini-exon (Comeau *et al.*, 1986; De Lange *et al.*, 1983; Nelson *et al.*, 1983). These are associated with unique metabolic and genetic functions, respectively. Other genes have particular organisations, like the large sub-unit (LSU) of the rRNA genes, which are present in several copies and separated by small spacers (Campbell *et al.*, 1987; Spencer *et al.*, 1987).

Most conserved genes across eukaryotes are also present in *Leishmania*, including for glucose and amino acid metabolism, some of which have been used for isoenzyme analysis. Future phylogenetic analysis may also take advantage of these universal genes.

1.2.5.4. Population genetics

Leishmania population structure is thought and considered to be essentially clonal (Tibayrenc and Ayala, 1991). The main arguments are the presence of over-represented identical phenotypes, linkage desequilibrium and absence of recombinant genotypes (Tibayrenc and Ayala, 1991; Tibayrenc *et al.*, 1990). The first two arguments apply to all *Leishmania* species, but some hybrids have been found between *L. major* and *L. arabica* (Evans *et al.*, 1987; Kelly *et al.*, 1991), *L. braziliensis* and *L. guyanensis* (Delgado *et al.*, 1997), *L. braziliensis* and *L. peruviana* (Dujardin *et al.*, 1995a) and *L. panamensis* and *L. braziliensis* (Belli *et al.*, 1994). In the isoenzyme characterization of *L. donovani sensu lato* (*s.l.*) by LeBlancq and Peters, (1986) a zymodeme (LON 50) is described which seems to be an ASAT hybrid between two otherwise identical zymodemes (LON 46 and 48). Hybrid formation seems to be more frequent in the *Viannia* subgenus, despite findings of mixed infections of other *Leishmania* species. Mixed *L. mexicana* / *L. braziliensis* infections were found in a sandfly (Barrios *et al.*, 1994), and human patients (Hernandez Montes *et al.*, 1998; Silveira *et al.*, 1984), as well as *L. braziliensis* and *L. donovani* (Oliveira Neto *et al.*, 1986b), in this case in different organs. A mixed *L. donovani* / *L. major* infection was also identified in a human patient and in the same organ (spleen) (Mebrahtu *et al.*, 1991), and two different *L. infantum* zymodemes were found in the same organs in a dog (Pratlong *et al.*, 1989). No evidence of hybrid formation was found, however,

even in sandfly experiments with *L. major* (Panton *et al.*, 1991), despite previous observation of promastigote fusion (Lanotte and Rioux, 1990). There is weak evidence that recombination, if occurring, might be a vertebrate host event (Kreutzer *et al.*, 1994) with fusion of amastigotes followed by meiosis. However, the duplication of DNA seen might be explained by duplication of DNA during mitosis.

It seems that *Leishmania* are capable of some form of recombination, even if not in the true sexual sense, but that population structure is clonal. The model of epidemic clonality (Maynard Smith *et al.*, 1993) may apply to *Leishmania* (Tibayrenc, 1998). In this model, the species is essentially sexual with clonal propagation in epidemic situations. *Leishmania* may be a special case of this model, capable of recombination, but essentially with clonal spread. A few clones may be amplified preferentially in epidemic situations and thus become dominant in the population. *Leishmania,* as clonal organisms, are capable of colonizing new environments with a single founder strain and initiating an epidemic. Furthermore, the likelihood of two simultaneous infections may be small due to host / vector requirements, but also to low transmission and incidence.

Studies on population structure of *Leishmania* are often biased because samples are often not sympatric (Tibayrenc, 1998). Methods used to detect genetic variability may not be sensitive enough to detect recombination events between closely related strains, which are the most likely to be able to cross.

1.3. The *Leishmania donovani* complex

1.3.1. Definitions of the complex

The parasites responsible for VL are those of the *L. donovani* complex. It was from VL cases that characteristic organisms were observed by Leishman and Donovan in 1903 (Jarry, 1999). Laveran and Mesnil propose the name *Piroplasma donovani* but later that year Ross creates the genus *Leishmania* for the new organisms. *Leishmania donovani* was thus the first to be described and has become the type species for the complex and for the genus *Leishmania*. In 1908 the parasite responsible for Mediterranean VL is named *L. infantum* by Nicolle, for the disease seems to be associated with infants. Castellani and Chalmers named the parasites causing VL in Sudan as *L. archibaldi* in 1919, which was later defined as zymodeme MON 82 (Rioux *et al.*, 1990). After VL was identified in South America by Chagas, in 1911, the parasite was finally named *L. chagasi* by Cunha and Chagas in 1937.

As more fully discussed in the next chapter, it is rather difficult to define species in *Leishmania*. Rioux *et al.* (1990) have used phenetic complexes of strains

instead, that would apply to monophyletic and closely related groups of zymodemes usually with particular epidemiological characteristics. The *L. donovani* complex, defined in that way, would include the designated species of *L. donovani*, *L. infantum*, *L. chagasi* and *L. archibaldi*. The taxonomic status of *L. archibaldi* is uncertain, as it has only been defined as zymodeme MON 82 and no epidemiological or clinical support has been found yet. The status of *L. chagasi* is also uncertain as most authors propose synonymy with *L. infantum*.

Several phylogenetic analyses, using either isoenzymes (Le Blancq *et al.* 1986; Mebrahtu *et al.* 1992; Rioux *et al.* 1990), DNA sequences (Piarroux *et al.* 1995), random amplified polymorphic DNA (RAPD) (Schonian *et al.* 1996), have shown that the *L. donovani* complex is monophyletic. The named species have been much more difficult to be identified as genetically monophyletic. Most analyses, however, have only used a small and often unbalanced number of strains of each group or used a limited number of characters which are difficult to apply for phylogenetics, such as isoenzymes, or single RAPD primers. Thus, except for isoenzyme typing, genetic diversity within the *L. donovani* complex has been poorly studied, despite the need of strain markers for diagnostics and better understanding of epidemiology.

More specifically, Indian VL does not seem to be identical to African VL. Despite clinical similarities, humans are the reservoir in Indian VL, whilst the main reservoir in Africa is essentially unknown. Furthermore, Indian *L. donovani* are isoenzymatically very homogenous whilst African *L. donovani* are much more diverse (Le Blancq *et al.* 1986; Rioux *et al.* 1990). Complicating the picture, Chinese strains were found to have a specific zymodeme (MON 35), which was related to an African rather than Indian zymodeme (Rioux *et al.*, 1990).

1.3.2. *Leishmania infantum* and *Leishmania chagasi* - Twins, sisters or cousins?

Cunha and Chagas in 1937 described the organism responsible for VL in the NW and named it *L. chagasi*, but Cunha (1938) later suspected that this organism and *L. infantum* were indistinguishable. Lainson and Shaw (1999), however, still prefer to maintain the species status for the agent of VL in the NW, defended by Travi *et al.* (1998) whilst other authors proposed synonymy (Grimaldi and Tesh, 1993; Rioux *et al.*, 1990).

A few genetic and phenetic methods have been reported to distinguish between *L. infantum* and *L. chagasi*, such as radiorespirometry (Decker-Jackson and

Tang, 1982), monoclonal antibodies (Santoro *et al.*, 1986), glycoconjugate ligands for promastigote internalization into murine macrophages (Palatnik *et al.*, 1990) and DNA fingerprinting (Ellis and Crampton, 1991), but all studied *L. chagasi* strains have isoenzyme profiles similar to the reference *L. infantum* strain IPT-1 (Cupolillo *et al.*, 1994; Momen *et al.*, 1987; Moreno *et al.*, 1984), the most common *L. infantum* zymodeme in Europe, and *L. chagasi* could not be distinguished from *L. infantum* by most genetic methods, such as restriction fragment length polymorphism (RFLP) (Beverley *et al.*, 1987), DNA probes (van Eys *et al.*, 1991) or RAPD (Schonian *et al.*, 1996).

The agent of visceral leishmaniasis in the Americas is *L. chagasi*, but all other species of the *L. donovani* complex belong to the Old World, which indicates an origin in the OW. By way of explanation, Lainson and Shaw propose that *L. chagasi* has been present for a long time in the American continent. The finding of healthy infected foxes (Silveira *et al.*, 1982) in the New Word, the existence of other New World reservoir hosts, such as sloths (Travi *et al.*, 1994), adaptation of the parasite to at least two vectors in the New World and wide geographical range suggest an ancient association (Lainson *et al.*, 1987), estimated at 2-3 million years (My) ago with the arrival of wild canids in the New World. However, *L. infantum* was also estimated to have diverged from *L. donovani* (Moreno *et al.*, 1984) 2My ago, and most studies, showed reduced diversity between *L. infantum* and *L. chagasi*, suggesting that, on the contrary, separation was quite recent.

1.4. Identification and phylogenies

At this point, it is necessary to define concepts and terminology. Taxonomy is the science of classification, which is the construction of taxonomic groups. Identification is the act of assigning an organism to an already established group. Phylogeny is the evolutionary history of organisms.

1.4.1. Traditional methods of identification and phylogenetic analysis of *Leishmania*

Before the advent of DNA technology, criteria used to characterize and classify *Leishmania* can be divided in three main groups: clinical and biological, immunological and biochemical.

1.4.1.1. Clinical and biological criteria

Since identification and isolation of *L. donovani* in 1903, and until the decade of 1950, other *Leishmania* species were predominantly named and classified according to geographic and clinical criteria (Pratlong and Lanotte, 1999). Thus, e.g. the names *L. aethiopica*, *L. mexicana*, *L. braziliensis*, were assigned, but also *L. infantum*, which was applied for mainly infantile kala-azar. Organisms isolated would be classified by the clinicians according to pathological presentation. It has, however, been recognized since that different species may produce similar clinical presentations and vice versa.

Although higher categories of protozoa have quite distinct morphologies, closely related species are often very similar and *Leishmania* are no exception. Despite morphological variation within the life cycle, *Leishmania* species are virtually indistinguishable morphologically. Only *L. major* and *L. braziliensis* have been distinguished by size from, respectively, *L. tropica* and *L. mexicana*. Even electron microscopy could not define species specific morphological differences, although microtubules have been used for differenciation (Pratlong *et al.* 1999).

As with bacteria, some researchers tried to characterize *Leishmania* using metabolic characteristics and growth rate. Unfortunately, *Leishmania* do not grow well in defined medium and require blood or serum, thus it was difficult to use metabolism as a character. Lainson and Shaw (1972) were able to distinguish between the faster growth of *L. mexicana* and the slower growth of *L. braziliensis* in Novy, McNeal and Nicolle (NNN) medium.

Characterization by means of animal infection criteria has been done by tropism and speed of development by Lainson and Shaw (1972) in the New World and Pratlong *et al.* (1986) in the Old World.

Lainson and Shaw (1979) were able to divide the genus *Leishmania* into two sub-genera by development in the vector gut: species of the *L. braziliensis* complex are also present in the posterior gut, which came to define the sub-genus *Viannia*, all other species belong to the sub-genus *Leishmania*.

1.4.1.2. Immunological criteria

A test using direct agglutination with live parasites (Noguchi-Adler) was able to distinguish between *Leishmania* species complexes but standardization was difficult to accomplish (Le Fichoux *et al.*, 1999). Schnur *et al.* (1972) used the excreted factor (phosphoglycans released by promastigotes and amastigotes during *in vitro* growth) for characterization. Phosphoglycans precipitate in presence of homologous rabbit

antiserum and are discriminated by Ouchterlony double diffusion. This technique differentiates between the main *Leishmania* species complexes. In 1981, a monoclonal antibody method was developed (McMahon Pratt and David, 1981) and differentiated between *Leishmania* species and even sub-species, but remained restricted to a few laboratories (Le Fichoux *et al.*, 1999).

1.4.1.3. Biochemical criteria

An heir to metabolic characterization, a technique, based on consumption of ^{14}C labelled amino acids and hydrocarbon molecules and named radiorespirometry, was developed by Decker-Jackson *et al.* (1977). Although a degree of correlation was found between metabolic profiles and clinical type, and it was claimed that radiorespirometry could distinguish between *L. infantum* and *L. chagasi*, characterization was not exact and it is technically demanding. Radiorespirometry never became adopted for identification purposes.

Lectins, proteins which are capable of specific recognition of sugars, were used in an agglutination test to distinguish between Old World *Leishmania* species (Schottelius, 1982) and even OW from NW species if enough lectins were used (Pratlong *et al.*, 1999). Many results were not convincing and lectin agglutination requires rigorous experimental conditions; it was also abandoned.

A biochemical technique which proved to be of enormous value, not only in *Leishmania*, but in many other organisms, is isoenzyme electrophoresis analysis (IEA). IEA was first used by Gardener *et al.* (1974) for *Leishmania* taxonomy. The electrophoretic mobility of malate dehydrogenase (MDH) differentiated between *Leishmania* groups, and many other enzymes were subsequently included. IEA is now the reference method for *Leishmania* identification and taxonomy, because of the adoption of standardization, and because it combines both specificity and stability of electromorphs with a high degree of intra-specific polymorphisms. IEA is amenable to typing by the enzyme profile, and to numerical taxonomy methods and cladistics, and thus suitable for phylogenetic studies. IEA was used in the present work and is described more fully in the introduction to Chapter 3.

Very recently, a method has been developed to identify the repeating phosphosaccharide units of *Leishmania* lipophosphoglycans using electrospray mass-spectrometry (Wilson *et al.*, 1999), which was used to differentiate *L. mexicana* from *L. major*. This approach is suitable for quick-typing of lipophosphoglycan repeats and was shown to detect alterations in repeat side chains caused by: culturing *L. major* promastigotes in the presence of L-fucose and *in vitro* metacyclogenesis of *L.*

major promastigotes. The authors anticipate that the method will be applicable to small samples of cultured field isolates or genetically-manipulated strains.

1.4.2. Molecular methods of identification and phylogenetic analysis

The first DNA based phylogenetic studies of *Leishmania* relied on the degree of cross-hybridization between genomic DNA of different species to extrapolate the degree of relatedness. With the discovery of DNA endonucleases (see introduction for Chapter 5) DNA hybridization was taken a step further to comparison of DNA fragments with specific probes. Band polymorphisms generated in this manner were called restriction fragment length polymorphisms (RFLPs) in what is called fingerprinting (Beverley *et al.*, 1984; el Hassan *et al.*, 1995; Ghalib *et al.*, 1992; Guizani *et al.*, 1994; Hu *et al.*, 1992; Lu and Hu, 1990; Macedo *et al.*, 1992; Massamba and Mutinga, 1992; Mendoza Leon *et al.*, 1995; Oskam *et al.*, 1998; Pascale *et al.*, 1992; Ramirez and Guevara, 1987; van Eys *et al.*, 1991; van Eys *et al.*, 1989). RFLP have been used for classification as well as for identification. Simple probe hybridization to genomic DNA has been a tool for sensitive diagnosis (Agatsuma, 1992; Barrios *et al.*, 1994; Benavides *et al.*, 1993; Blackwell, 1992; Briones *et al.*, 1992; Esseghir *et al.*, 1993; Howard *et al.*, 1991; Hu *et al.*, 1992; Laskay *et al.*, 1991; Lu and Hu, 1990; Maingon *et al.*, 1993; Massamba and Mutinga, 1992; Singh, 1997; Wilson, 1995). Applicability, in terms of specificity, resolution and sensitivity, of RFLP or fingerprinting to taxonomic groups depends on the probe used. However, the RFLP method is not very flexible because it is dependent on the probe used and only detects variation in the size of fragments harbouring conserved sequences.

In a variation of the method, but without restriction, it was possible to study chromosomes and their gene organization. As mentioned before (see 1.2.5.1.) PFGE is used to separate and thus visualize *Leishmania* chromosomes. By use of probes, homologous chromosomes can be identified (Wincker *et al.*, 1997) and karyotypic analysis has been used for comparison between *Leishmania* strains and species (Dujardin *et al.*, 1995b; Dujardin *et al.*, 1995c; Giannini *et al.*, 1986; Soto *et al.*, 1995).

With PCR, it is possible to amplify unstable / more polymorphic regions using conserved flanking regions for primer hybridization, such as intergenic regions (Cupolillo *et al.*, 1995; Eisenberger and Jaffe, 1999; Hassan *et al.*, 1993), kinetoplast DNA (Breniere *et al.*, 1999), or satellite DNA (Ravel *et al.*, 1995; Rossi *et al.*, 1994; Russell *et al.*, 1999). PCR methods, in their many experimental modifications, are suitable for studying different aspects of genomes and can be applied to different aspects of taxonomy. Band size polymorphisms can be generated by variation in

number of repeated nucleotides, such as mini or microsatellites (Ravel *et al.*, 1995; Rossi *et al.*, 1994; Russell *et al.*, 1999), or by random hybridization of primers within the genome (Banuls *et al.*, 1999a; Banuls *et al.*, 1999b; Bhattacharyya *et al.*, 1993; Eisenberger and Jaffe, 1999; Garcia *et al.*, 1998; Gomes *et al.*, 1995; Maingon *et al.*, 1993; Motazedian *et al.*, 1996; Noyes *et al.*, 1996; Oskam *et al.*, 1998; Schonian *et al.*, 1996; Singh, 1997; Welsh and McClelland, 1990; Williams *et al.*, 1990), as in RAPD (Williams *et al.*, 1990) or arbitrarily primed - PCR (AP-PCR) (Welsh and McClelland, 1990). Amplified products can be characterized with probes (Breniere *et al.*, 1999; Degrave *et al.*, 1994; Guevara *et al.*, 1992; Maingon *et al.*, 1993; Qiao *et al.*, 1995; Rodgers *et al.*, 1990; Rodriguez *et al.*, 1994; Wilson, 1995), and detected by radioactivity, chemioluminescence, or be linked to an ELISA (PCR- solution hybridization enzyme-linked immunoassay, PCR-SHELA) (Qiao *et al.*, 1995). Amplification products can be further characterized by restriction analysis (PCR-RFLP) (Singh, 1997; Victoir *et al.*, 1998). Multiplex PCR can be used with or without nested-PCR to generate differently sized products in the same reaction (Belli *et al.*, 1998; Harris *et al.*, 1998).

A wide range of other techniques has been developed, which could be applied in PCR based approaches to *Leishmania* taxonomy. Separation of fragments in denaturing conditions (Stothard *et al.*, 1997) by denaturing gradient gel electrophoresis (DGGE, Cariello *et al.*, 1988), single stranded conformational polymorphism (SSCP, Orita *et al.*, 1989), temperature gradient gel electrophoresis (TGGE, Riesner *et al.*, 1989), which enables distinction between different sequences with identical lengths. Specific amplification of restriction products, by amplified fragment length polymorphism (AFLP) (Vos *et al.*, 1995). Single nucleotide polymorphisms can be detected in PCR reactions by molecular beacons (Marras *et al.*, 1999; Tyagi *et al.*, 1998; Tyagi and Kramer, 1996), which are hair pin probes that fluoresce upon specific hybridization, or by TaqMan technology (Heid *et al.*, 1996; Holloway *et al.*, 1999) (probes that fluoresce when a chelator is released from the annealed probe by the DNA polymerase in a PCR reaction. Micro or mini oligonucleotide arrays (Chetverin and Kramer, 1993; Ramsay, 1998) can be used in fast DNA sequencing (Chetverin and Kramer, 1993), to detect mutations (Ravine, 1999) or to study phylogenetic relationships (Hacia *et al.*, 1998). The number of new techniques is increasing and mostly driving towards higher processivity, lower cost and applicability to a wider range of problems, although many may never become established.

The most powerful technique for phylogenetic analysis is DNA sequencing (see introduction to Chapter 6), which greatly benefited from the development of PCR. DNA sequencing is invaluable for detailed analysis of genetic diversity and, thanks to ever more cheap and easy methods, may even be used for identification.

1.4.3. Phylogenetic reconstruction

The first classifications of *Leishmania* were of the Linnean type, based on a small number of mostly extrinsic characters, which by exclusion would characterize taxonomic groups at any level. Categories were essentially applied for convenience of arrangement. Lainson and Shaw (1987) proposed the most complete classification of the genus.

When it was recognized that all living organisms are linked by descent from common ancestors, in other words, that species evolve from others, it became evident that classifications should be based on evolutionary relationships rather than similarities or convenience for the scientist. It thus became necessary to devise methods for establishing degrees of relationship and speciation events.

Initially, evolutionary classifications were dependent on the expertise of the scientists. Expert knowledge, based on years of studying organisms, frequently only morphological characters, was necessary for classifications. Frequently, experts would disagree, and a more scientific and reliable measure for evolutionary relationships became necessary.

Numerical taxonomy (Sokal and Sneath, 1963), assumed that calculating distances from a large number of characters, pooled and without hierarchy, would provide a reliable measure of how related organisms were, without any evolutionary hypothesis. Resulting classifications are called Adansonian or phenetic, for they are based upon phenetic characters. This method became more practical thanks to the development of computational facilities. The use of numerical taxonomy with *Leishmania* was applied to biochemical, such as isoenzyme profiles, and now also molecular characters, DNA sequences, RFLPs and RAPDs. Demes or identical profiles represent operational taxonomic units (OTUs) and strains can be thus identified.

Almost at the same time, critics claimed that evolution can only be deciphered by analysis of shared derived homologous characters (synapomorphisms) and thus only cladistic classifications (Hennig, 1965), based on clades or monophyletic groups or organisms are true. Although cladistics seems a more realistic methodology, choice of adequate characters (often few) and distinction of homologies from homoplasies is often still a matter of the researcher's judgement.

Traditional characters on which to base phylogenies were replaced by protein and enzyme electrophoresis and later by DNA methods, which probe the very core of evolutionary change. The several methods that are or were used to produce phylogenies and classifications have already been described in the section above.

Depending on the methodology and concepts used, several models were devised for determining phylogenies. Broadly, the methods can be divided into distance methods, which are essentially phenetic, character methods, like parsimony, which are cladistic in nature, and maximum likelihood. This separation of methods is not straightforward, however, because given enough data, parsimony becomes almost phenetic in the sense that contradictory information has to be weighted and phenetic methodology can produce effectively cladistic trees because true phenetic similarity implies evolutionary relatedness.

1.4.4. Methods for construction of phylogenies

Character based methods produce phylogenetic trees through search or optimality criteria (Page and Holmes, 1998) with explicit functions to relate tree and data (or model of evolution). The quality of the tree is thus evaluated. The foremost search methods are maximum likelihood and maximum parsimony. Both methods perform better with small numbers of taxa, and a limitation is the required computing time, most notably for maximum likelihood. Maximum likelihood produces a tree that makes the data the most probable evolutionary outcome (Page and Holmes, 1998). It is thus a probabilistic model (Weir, 1996) which assumes an evolutionary model. The model assumes that changes are more likely along long branches and so, branch lengths are important (Weir, 1996). All data are used in this method. Maximum parsimony searches for trees with minimum total lengths and can thus produce several alternative trees. Only informative sites are used (Weir, 1996) and parsimony relies on few assumptions (Page and Holmes, 1998), namely: minimum total length of tree and maximum similarity between taxa. Because branch lengths are not important, parsimony is sensitive to the long branch attraction phenomenon (Page and Holmes, 1998; Weir, 1996): lineages in which more evolutionary change occurred are placed nearer to the root. This method performs better with more data and small change. There are a number of algorithms for maximum parsimony search. Wagner (Farris, 1970; Kluge and Farris, 1969) assumes reversiblility of character states. Trees can thus be rooted at any node without change in tree length (Weir, 1996). Dollo parsimony (Farris, 1977) weights probabilities of reversion and is thus the method of choice for restriction fragments. Polymorphism parsimony assumes

retention of polymorphisms and Camin-Sokal (1965) considers that change to ancestral states is irreversible (Avise, 1994).

Distance methods use a measure of similarity on which the quality of the resulting tree is dependent (Weir, 1996). These are clustering methods, which produce a single tree, and are fast but results can sometimes depend on order of addition of taxa (Page and Holmes, 1998). The unweighted pair-group method with arithmetic mean (UPGMA) (Sneath & Sokal, 1973) produces a rooted tree because an equal rate of evolution (molecular clock) is assumed for all lineages (Avise, 1994; Page and Holmes, 1998; Weir, 1996). Despite these assumptions it was found to perform well (Avise, 1994). The neighbour-joining method (Saitou and Nei, 1987) is the most used and produces an unrooted tree by determining the closest pairs to minimize the length of the tree (Weir, 1996). The resulting tree is correct for purely additive data (Weir, 1996) and performs as the minimum evolution method (a search method for distance data) (Page and Holmes, 1998) if with unequal rates of evolution (Avise, 1994). Another method is Fitch and Margoliash (1967), an additive approach which uses a weighted least squares algorithm (Weir, 1996) and produces an unrooted tree. The optimal tree is searched for by percent standard deviation. It performs better than UPGMA if rates of change are unequal in different lineages.

Distance methods assume that distances are metric (Page and Holmes, 1998), or:

1. are not negative
2. are symmetric (the distance between taxa a and b is the same as b and a)
3. have triangle inequality (the distance between a and c is smaller than the sum of the distance between a and b and between b and c)
4. are distinct (or that a is not identical to b).

Distances must also be ultrametric, or

5. the two largest distances must be equal, and thus must arise through an equal rate of evolution.

Finally, distances must be additive, or

6. the sum of the distance between a and b and between c and d, must be smaller or equal to the sum of the distances between a and c and between b and d, or the sum of the distances between a and d and between b and c.

One way of evaluating compliance with ultrametric requirements is to compare the difference between the distances calculated from the tree and the observed distances. This is done by a measure of cophenetic correlation.

To correctly measure genetic distance a few assumptions are made (Page and Holmes, 1998):

- all nucleotide sites change independently,
- substitution rate is constant,
- base composition is at equilibrium,
- probabilities of substitutions are the same for all sites at all times.

These assumptions are often violated by genetic data. For example, regions coding for stems in rRNA genes are strongly dependent on the complementary strand (Wheeler *et al.*, 1988). The possible mutable sites made available by selection may be so low, that even fast evolving sequences can show less divergence than slower ones (Page and Holmes, 1998). In most organisms the distribution of mutation rates follows a Gamma distribution with $\alpha \leq 1$, and thus exhibits a L shaped curve, with only a small number of sites exhibiting a high number of polymorphisms, and the large majority with only a few polymorphisms. In saturated sequences, mutations can often reverse characters to ancestral states, thus confusing phylogenies. Substitution saturation can be detected by comparison of branch lengths from parsimony analyses with pairwise distances.

Except for maximum likelihood, most methods do not intrinsically provide a measure of the reliability of the tree. There are, however, resampling approaches which allow evaluation of the dependence of the tree on the sampled data (Weir, 1996). The most used method is bootstrapping, which produces a new data set by randomly selecting characters from the original data set (Weir, 1996). Resampling can include duplication of some characters. Another, less used method, Jacknife, produces new data sets with (n-1) characters. Usually 100 or 1000 new data sets are used to produce the same number of alternative trees from which a majority rule consensus tree can be produced, which will include the most common nodes of all other trees. It can be said that a branch which is present in 95% of the trees can be trusted within a 5% confidence interval. Examples of other consensus trees are strict consensus, which only includes shared groups or splits, and Adam's consensus, which maintains the groups found in all trees (Page and Holmes, 1998).

To choose an adequate method for reconstructing pylogenies, several properties have to be considered (Page and Holmes, 1998), particularly when analysing large data sets. Efficiency, or how fast a method is, is very important. UPGMA is very fast, and among search criteria methods, parsimony is much faster than likelihood, and can thus search more trees. The power of the method is a measure of how many data are necessary to produce a reliable result and consistency evaluates if a method will converge on the right tree if more data are added. A robust method is less sensitive to violation if it indicates whether assumptions are being violated.

Maximum likelihood is often considered, but not consensually (Steel and Penny, 2000), the best method because it combines a cladistic approach with branch lengths, it is a very consistent method and the only method that intrinsically produces a measure of falsifiability, or violations of assumptions. Being such a thorough method, maximum likelihood is also very slow and not at all efficient with large data sets. Neighbour-joining can be just as good in practice but more efficient.

1.5. Objectives

The present thesis describes the work done to achieve the following objectives:

- to study the genetic diversity in the *L. donovani* complex;
- to determine whether *L. chagasi* and *L. infantum* are valid species and their position in the *L. donovani* complex;
- to determine if the antigen GP63 is conserved between *L. infantum* and *L. chagasi* strains;
- to characterise genetically Portuguese strains of *L. infantum*.

2. Materials and Methods

2.1. Parasites

2.1.1. *Leishmania* strains

A group of strains representing the *L. donovani* complex and belonging to its four named species was assembled for assessment of taxonomic relationships (Fig. 7). A collection of reference strains (RS) of *L. major, L. tropica* and *L. aethiopica* and DNA from New World *Leishmania* species, provided by Debbie Nolder were used as outgroups (Table II).

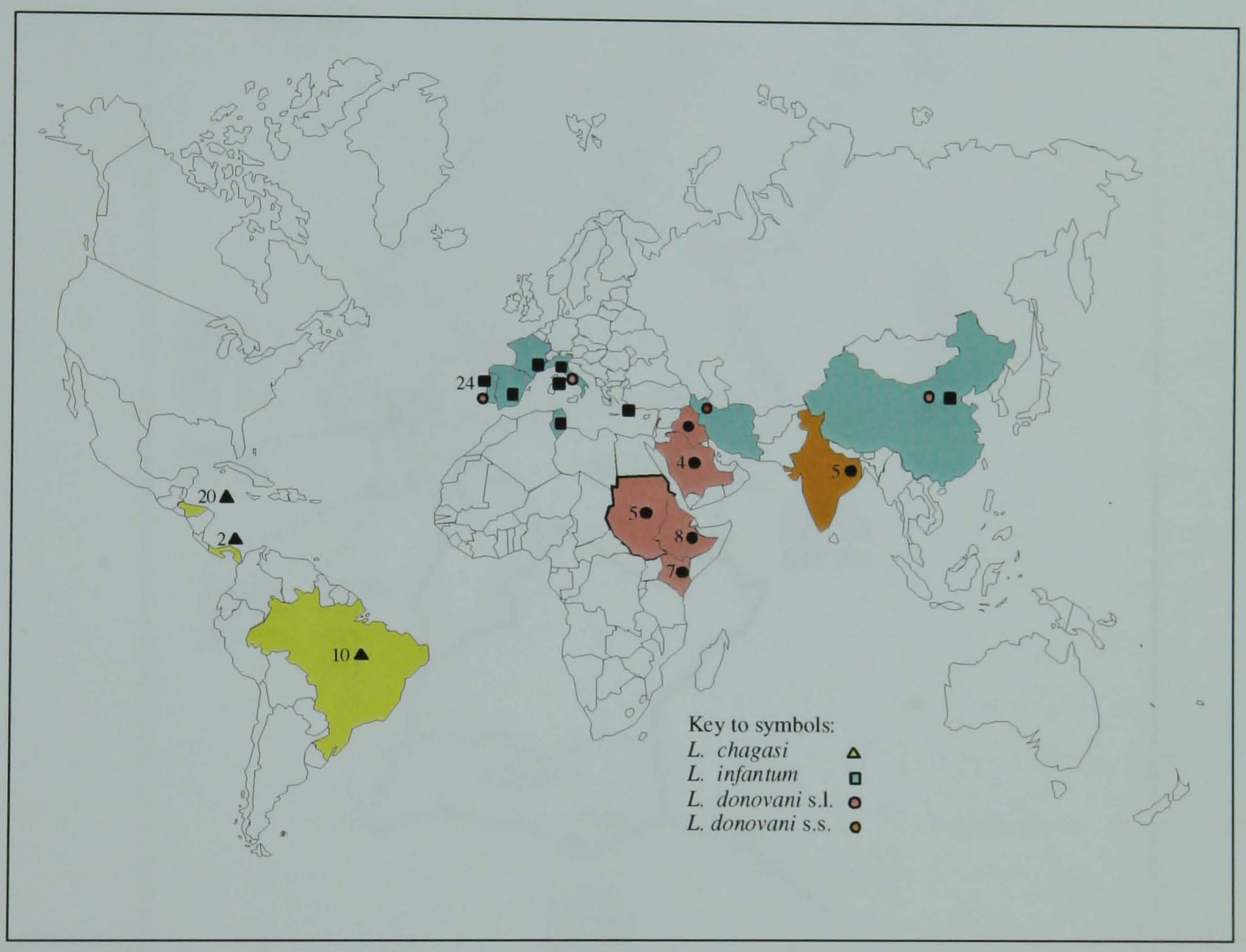


Figure 7 - Countries of origin of the *Leishmania donovani* complex strains. Some *L. donovani* strains (shown in open circles) were isolated in countries endemic for *L. infantum*.

A group of 13 *L. chagasi* strains was assembled (Table III), nine from the LSHTM cryobank and four from the Belem cryobank, Brazil. Two strains were isolated in Panama from human patients and the remainder from Brazil. Three strains were isolated from naturally infected dogs in Teresina, an important Brazilian focus where VL is epidemic, and four strains (Table III) were from Amazonian foxes, *Cerdocyon thous*. The remainder were chosen at random from the LSHTM cryobank, and

included canine and human cases. Dr. Harry Noyes provided DNA from 20 *L. chagasi* strains from VL or CL human patients of Honduras, three of which were used here.

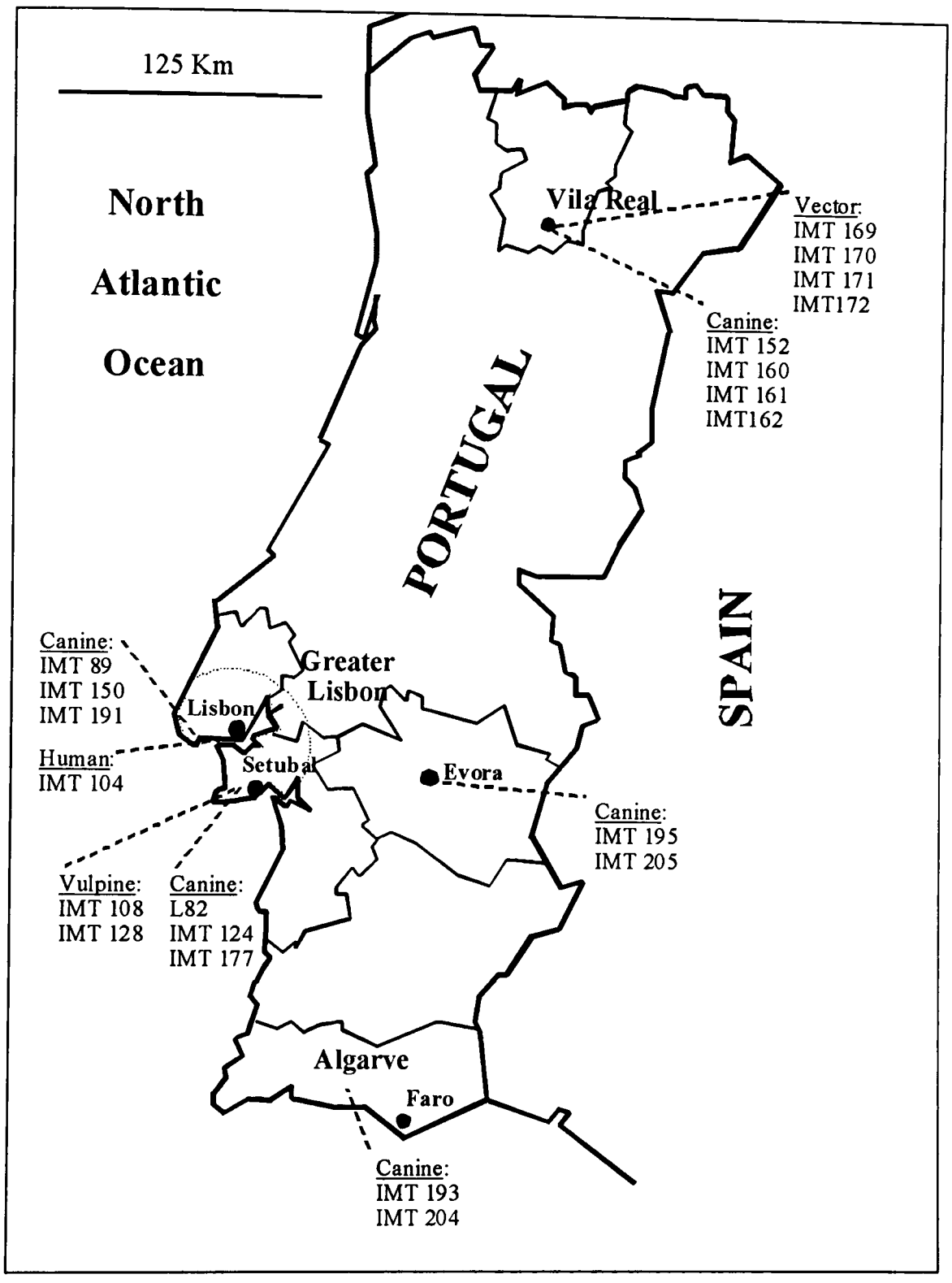


Figure 8 - Map of Portugal, showing the main endemic areas and the origin of the strains used in this project.

A total of 31 *L. infantum* strains were selected (Table IV). Twenty two were isolated in Portugal, and nine in other countries (China, Cyprus, France, Italy, Malta, Spain, Sudan, Tunisia). A CL isolate from Cyprus and a MON 34 from China were included as representatives of the diversity within *L. infantum*. Representative strains from the three main Portuguese foci (North, Lisbon and Algarve) and also from the Évora region that has recently been indicated as being another focus (Semião Santos *et al.*, 1995), Fig. 8, were chosen. Four canine isolates and four vector

isolates were from the North (Table IV), two of the latter were zymodeme MON 24. The human isolate was from the Lisbon region, and of the six canine isolates, three were from the second endemic area of Setubal, the origin of the two fox isolates (Table IV). The Algarve region contributed two canine isolates (Table IV).

L. donovani strains (Table V) were selected from the extent of the distribution (Fig. 7) and retrieved from the LSHTM cryobank. The Portuguese strain, IMT 180, was isolated from a HIV patient and is thought to have been imported. The other Portuguese strain had been isolated from *Sergentomyia minuta*, usually considered as vector of lizard *Leishmania.* Two different stabilates were taken from the cryopreserved stock and transported to the UK at different times. One strain of *L. archibaldi* (Gebre 1) was also included. At least one strain (Gilani) had been typed as *L. infantum* in the Montpellier system, because of its aspartate aminotransferase (ASAT, or GOT) profile, but as *L. donovani s.l.* by Le Blancq (1986).

2.1.2. Cultures

Ampoules with cryopreserved *Leishmania* cultures from the LSHTM cryobank were retrieved from liquid nitrogen and quickly thawed in water, then immediately transferred to 0.4ml of proline balanced salts solution (PBSS) and inoculated into Novy, McNeal and Nicolle (NNN) blood slopes (Evans *et al.*, 1989): 1.4% w/v purified agar (Oxoid), 0.9% weight / volume (w/v) sodium chloride (NaCl), in distilled water, autoclaved and supplemented with 10% v/v defibrinated rabbit blood. Strains from other cryobanks were immediately transferred upon arrival to fresh NNN slopes with PBSS overlay: 5.37mM potassium chloride (KCl), 167μM di-sodium hydrogen phosphate dodecahydrate ($Na_2HPO_4.12H_2O$), 441μM potassium dihydrogen phosphate (KH_2PO_4), 1.26mM calcium chloride dihydrate ($CaCl_2.2H_2O$), 0.41mM magnesium sulphate ($MgSO_4.7H_2O$), 0.49mM magnesium chloride ($MgCl_2.6H_2O$), 136.89mM sodium chloride (NaCl), 8.69mM L-proline, 2.82μM phenol red, in distilled water and adjusted to pH 7.2 with solid Tris (Tris (hydroxymethyl) aminomethane). The parasites were initially subcultured in a small amount of liquid overlay (~0.5ml) and, if actively growing, enough medium was added to cover the slope. Actively growing (3^{rd} or 4^{th} day) rich cultures were inoculated in a proportion of 1:5, inoculum to fresh medium into suplemented α-Minimum Essential Medium (S-α-MEM) (Kar *et al.*, 1990): minimum essential medium (MEM) Eagle, alpha modification powder for 1 litre, 0.11% w/v sodium hydrogen carbonate ($NaHCO_3$), pH 7.5, 4mM L-glutamine, 16.6mM D-glucose, 11.3μM folic acid, 8.2μM D-biotin, 7.7μM haemin, 38.4mM HEPES (N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulforic acid, sodium salt) and 0.296mM adenine in distilled water and 10% volume / volume (v/v) (to page 69)

Table II - Outgroup *Leishmania* strains.

	Species WHO code	Zymodeme LON	Zymodeme MON	Cryobank of origin	Techniques
	L. aethiopica				
	MHOM/ET/1970/L96	27	-	L	3, 4im
OLD	MHOM/ET/1972/L100 *	-	14	L	3
WORLD	*L. major*		4		
	MHOM/SU/1973/5-ASKH*	1		L	1, 3, 4im
	L. tropica				
	MHOM/SU/1974/K27*	12	60	L	3, 4im
	L. amazonensis				
	MHOM/BR/1973/M2269*†	-	-	L	-
	L. braziliensis				
	MHOM/BR/1984/LTB300*†	200[(1)]	-	L	-
	MHOM/BR/1975/M2903*†	211[(1)]	43	L	-
NEW	*L. guyanensis*				
	MHOM/BR/1975/M4147*†	-	-	L	-
WORLD	*L. lainsoni*				
	MHOM/BR/1981/M6426*†	-	-	L	-
	L. mexicana				
	MHOM/BZ/1982/BEL21*†	-	-	L	-
	L. panamensis				
	MHOM/PA/1971/LS94*†	-	-	L	-
	L. peruviana				
	MHOM/PE/1984/LC39*†	-	-	L	-

* - WHO reference strains. † - DNA only. [(1)] assigned by Deborah Nolder (PhD thesis). WHO codes: CSSS/cc/YY/N : host(**C**- class; **SSS** - species)/country/Year/ strain Name. Host: M - mammal; HOM - *Homo sapiens*. Country: BR: Brazil; ET: Ethiopia; PA: Panama; PE: Peru; SU: former Soviet Union. Cryobank: L - LSHTM. Techniques applied in this work: 1 - Lmet9; 3 - RAPD; 4 - RFLP (i - ITS, m - mini-exon).

Table III - *L. chagasi* strains

Number	WHO code	Patho-logy	Zymodeme		Cryobank of origin	Techniques
			LON	MON		
C	MHOM/BR/1974/PP75[a]	VL	49	1	L	1, 2, 3, 4
C3	MCAN/BR/1984/CO910	CVL	-	-	L	2a, 3, 4m
C6	MCAN/BR/1989/DOG124	CVL	-	-	L	2, 3, 4l
C7	MCAN/BR/1989/DOG136	CVL	-	-	L	1, 3
C8	MCAN/BR/1989/DOG118	CVL	-	-	L	1, 3
C13	MCER/BR/1981/M6445	VL	-	-	B	1, 2, 4
C9	MCER/BR/1983/M7633	VL	-	-	B	3
C12	MCER/BR/1989/M12084	VL	-	-	B	3c
C11	MCER/BR/1989/M12085	VL	-	-	B	3
C5	MHOM/BR/1984/M8270	-	-	-	L	3
C4	MHOM/BR/1985/M9702	-	-	-	L	3
C16	MHOM/HN/1988/HN115†	CL	-	-	LV	4i
C17	MHOM/HN/1988/HN122†	CL	-	-	LV	4
C20	MHOM/HN/1993/HN336†	CL	-	-	LV	4
C2	MHOM/PA/1978/WR285	-	-	-	L	1, 2, 3, 4
C1	MHOM/PA/1980/WR341	-	-	-	L	1

* - WHO reference strain. † - DNA only. WHO codes: CSSS/cc/YY/N : host(C- class; SSS - species)/country/Year/strain Name. Host: I - insect; M - mammal; HOM - *Homo sapiens*; CAN - *Canis familiaris*; CER - *Cerdocyon thous*; LUT - *Lutzomyia* sp.Country: BR: Brazil; HN: Honduras; PA: Panama. Cryobank: L - LSHTM; B - Belem; LV - Liverpool. Techniques applied in this work: 1 - isoenzyme analysis; 2 - *mspC*3' end sequence (2a - partial sequences); 3 - RAPD; 4 - RFLP (i - ITS, g - gp63, m - mini-exon).

Table IV - *L. infantum* strains.

Number	WHO code	Zymodeme		Cryobank	Region	Techniques
		LON	MON	of origin	**	
I	MHOM/TU/1980/IPT-1*	49	1	L	-	1, 2, 3, 4
I23	IARI/PT/1989/IMT 169	-	1	M	North	3
I24	IARI/PT/1989/IMT 170	-	1	M	North	2a, 3
I25	IARI/PT/1989/IMT 171	-	24	M	North	1, 2, 3, 4iml
I26	IARI/PT/1989/IMT 172	-	24	M	North	1, 2a, 3, 4icl
I3	MCAN/FR/1982/Pharoah	49	-	L	-	1, 2a, 3, 4
I1	MCAN/PT/1981/L82	49	-	L	Lisbon	1, 2a, 3, 4
I10	MCAN/PT/1981/IMT 89	49	-	L	Lisbon	1, 3
I11	MCAN/PT/1982/IMT 124	-	-	L	Lisbon	1, 3
I8	MCAN/PT/1987/IMT 150	-	1	L	Lisbon	3
I13	MCAN/PT/1993/IMT 193	-	1	I	Algarve	3
I9	MCAN/PT/1987/IMT 152	-	1	L	North	1, 3
I2	MCAN/PT/1988/Rebelo 2	-	-	L	Other	1, 2a, 3
I20	MCAN/PT/1989/IMT 160	-	1	M	North	3
I22	MCAN/PT/1989/IMT 161	-	1	M	North	3
I21	MCAN/PT/1989/IMT 162	-	1	M	North	3
I27	MCAN/PT/1991/IMT 177	-	1	M	Lisbon	3
I29	MCAN/PT/1993/IMT 191	-	1	M	Lisbon	3
I18	MCAN/PT/1994/IMT 204	-	1	I	Algarve	3
I19	MCAN/PT/1994/IMT 205	-	1	I	Other	3
I30	MCAN/PT/1994/IMT 195	-	1	M	Other	1, 3
I16	MHOM/CN/1980/Strain A	-	34	L	-	1, 2, 3, 4
I17	MHOM/CY/1963/L53	-	-	L	-	1, 2a, 3, 4
I4	MHOM/ES/1987/Lombardi	-	-	L	-	1, 2a, 3, 4iml
I5	MHOM/FR/1978/LEM75	-	1	L	-	2a, 3
I15	MHOM/IT/1981/Alessandro	-	1	L	-	1, 2a
I31	MHOM/MT/1985/Buck	49	78	L	-	1, 2, 3, 4icl
I6	MHOM/PT/1982/IMT 104	-	1	L	Lisbon	1, 3, 4c
I33	MHOM/SD/0000/CAM 1	-	-	C	-	2, 4icl
I12	MVUL/PT/1983/IMT 108	49	-	L	Lisbon	3
I7	MVUL/PT/1984/IMPT 128	-	1	L	Lisbon	3

* - WHO reference strain. ** - Portuguese regions of isolation (see Figure 2). WHO codes: CSSS/cc/YY/N : host (**C**- class; **SSS** - species)/**c**ountry/**Y**ear/strain Name. Host: I - insect; M - mammal; HOM - *Homo sapiens*; CAN - *Canis familiaris*; VUL - *Vulpes vulpes*; ARI - *Phlebotomus ariasi*. Country: CN: China; CY: Cyprus; FR: France; IT: Italy; MT: Malta; PT: Portugal; SP: Spain; TU: Tunisia. Cryobank: C - Cambridge; L - LSHTM; I - IHMT; M - Montpellier. Techniques applied in this work: 1 - Lmet9; 2 - *mspC*3' end sequence (2a - partial sequences); 3 - RAPD; 4 - RFLP (i - ITS, g - gp63, m - mini-exon).

Table V - *L. donovani* strains (according to the classification used by the LSHTM).

Number	WHO code	Zymodeme		Cryobank of origin	Techniques
		LON	MON		
D	MHOM/IN/1980/DD8*†	41	2	L	1, 2, 3, 4
D28	MHOM/ET/1972/Gebre 1**	50	82	L	1, 2, 3, 4
D21	IMAR/KE/1962/LRC-L57	44	37	L	1, 2, 3, 4
D35	ISER/PT/1993/IMT 188	-	-	I, Algarve	1, 2, 3, 4
D22	MARV/SD/1962/LRC-L64	48	-	L	1, 3, 4
D23	MCAN/IQ/1981/SUKKAR 2	43	-	L	1, 2a, 3, 4
D24	MCAN/IT/1976/DORA	50	-	L	1, 2, 3, 4
D25	MCAN/KE/0000/D2	45	-	L	1, 2, 3, 4
D26	MHOM/CN/0000/Wangjie-1	-	35	L	1, 3, 4
D13	MHOM/ET/0000/Hussen	42	-	L	1, 2, 3, 4
D12	MHOM/ET/0000/Ayele 5	52	-	L	1, 2, 3, 4
D27	MHOM/ET/0000/Ayele8	56	-	L	1, 2, 3, 4
D1	MHOM/ET/1967/HU3 (LV9)*	46	18	L	1, 2, 3, 4
D14	MHOM/ET/1982/Bekele	42	-	L	1, 3, 4
D15	MHOM/ET/1984/Addis 142	-	-	L	1, 3, 4
D29	MHOM/ET/1984/Addis 164	-	83	L	1, 2, 3, 4
D4	MHOM/IN/1982/Patna 1†	41	-	L	1, 2, 3, 4
D6	MHOM/IN/1977/Chowd-X†	-	-	L	1, 2, 3, 4
D7	MHOM/IN/1979/STL1-79†	-	-	L	1, 2, 3, 4
D8	MHOM/IN/1982/Nandi 1†	41	-	L	1, 3, 4
D30	MHOM/KE/0000/Neal-R1	56	-	L	1, 3, 4
D2	MHOM/KE/1967/MRC(L)3	-	-	L	1, 2, 3, 4
D3	MHOM/KE/1973/MRC74	51	-	L	1, 2, 3, 4
D31	MHOM/KE/1975/Mutinga H9	56	32	L	1, 2, 3, 4
D16	MHOM/KE/1980/Ndandu 4A	44	-	L	1, 2, 3, 4
D17	MHOM/LB/1984/Salti 4	-	-	L	1, 2, 3, 4
D34	MHOM/PT/1992/IMT 180	-	18	I	1, 2, 3, 4
D32	MHOM/SA/1981/Jeddah KA	42	31	L	1, 3, 4
D11	MHOM/SA/1987/VL6	-	-	L	1, 3, 4
D9	MHOM/SA/1987/VL23	-	-	L	1, 3, 4
D10	MHOM/SA/1987/VL29	-	-	L	1, 2, 3, 4
D18	MHOM/SD/0000/Khartoum	46	18	L	1, 2, 3, 4
D33	MHOM/SD/1982/GILANI	48	30	L	1, 2, 3, 4
D19	MHOM/SD/1985/A22	-	-	L	1, 2, 3, 4
D20	MHOM/SD/1987/UGX-marrow	-	-	L	1, 3, 4
D5	MMER/IR/1996/MESH-17	50	-	IRAN	1, 3, 4

* - WHO reference strain. ** - *L. archibaldi*. † - *L. donovani* s.s.

WHO codes: CSSS/cc/YY/N : host(**C**- class; **SSS** - species)/country/**Y**ear/strain Name. Host: I - insect; M - mammal; HOM - *Homo sapiens*; CAN - *Canis familiaris*; MAR - *Phlebotomus martini*; ARV - *Arvicanthis* sp; MER - *Meriones* sp; SER - *Sergentomyia minuta*. Country: CN - China; ET: Ethiopia; IN: India; IR - Iran; IQ - Iraq; IT: Italy; KE: Kenya; LB- Lebanon; PT: Portugal; SA - Saudi Arabia; SD - Sudan. (continues next page).

(continuation of Table V) Cryobank: L - LSHTM; I - IHMT, Lisbon, Portugal; M - Laboratoire d'Ecologie Medicale et Pathologie Parasitaire, Montpellier; B - Instituto Evandro Chagas, Belem, Brazil; LV - Liverpool School of Tropical Medicine and Hygiene, UK. Techniques applied in this work: 1 - Lmet9; 2 - *mspC*3' end sequence (2a - partial sequences); 3 - RAPD; 4 - RFLP (c - gp63 ITG/CS, i - ITS, l - gp63 ITG/L, m - mini-exon).

sterile heat sterile heat inactivated foetal calf serum, 50mg/l gentamicin. Cultures were subcultured into increasing volumes of liquid medium (at 1:5 or 1:10, inoculum to fresh medium) until enough parasites (10^6-10^7/ml) were obtained for DNA or enzyme extraction.

2.2. Enzyme extraction

Enzyme extractions from *Leishmania* were done as described by Evans *et al.* (1989). Parasites from 120ml late logarithmic phase cultures (usually day 5), were collected and centrifuged (700g for 20 min at 4°C) and the pellet was sequentially washed with 10ml and 1.0ml cold PBSS. An equal volume of 2mM stabiliser solution [2mM ethylenediaminetetraacetic acid disodium salt (EDTA), 2mM, ε-aminocaproic acid, 2mM dithiothreitol in distilled water] was added to the pellet. The suspension was treated with three cycles of freezing in liquid nitrogen and thawing, after which it was centrifuged at high speed (13000g) for 30 min at 4°C. The supernatant was collected, beaded in approximately 10µl droplets and stored in liquid nitrogen until needed.

2.3. DNA extraction

Extraction of genomic DNA was adapted from Kelly (1993). Parasites from a 15ml mid-logarithmic phase bulk culture were collected by centrifugation (700g for 20 min at 4°C) and washed three times in ice-cold sterile phosphate buffer saline (PBS), pH 7.2 (136.89mM sodium chloride (NaCl), 2.68mM potassium chloride (KCl), 12mM di-sodium hydrogen phosphate (Na_2HPO_4), 1.76mM potassium di-hydrogen phosphate (KH_2PO_4), in distilled water). The pellet was resuspended in sterile 1ml cell lysis buffer (CLB; 0.125M NaCl, 0.125M EDTA, 2.5% w/v sodium dodecyl phosphate (SDS), 0.125M Tris, in distilled water, pH 8.0) with 100µg/ml proteinase K, and incubated at 37°C overnight or 56°C for 3 hours.

The DNA was freed from contaminants in the lysate by sequential extraction with equal volumes of phenol, phenol/chloroform (1:1) and chloroform, each followed by a centrifugation step (6000g for 10min at 4°C). The DNA was precipitated from the aqueous phase with 2.5 volumes of ice cold absolute ethanol, transferred into 200-500µl of sterile distilled water using a glass pipette loop, and left to dissolve. The DNA extract was incubated at 37°C for 1 hour with 5ng of heat-treated ribonuclease A (Boehringer Mannheim). All DNA samples were stored at 4°C.

DNA concentration, purity and integrity were estimated by visualization on an agarose gel: 1% agarose in tris acetate EDTA (TAE) buffer (40mM tris, 1.142% v/v glacial acetic acid, 10% w/v 10mM EDTA (pH8.0) in distilled water) with 1µg/l

ethidium bromide (EtBr). The DNA concentration and purity were determined by spectrophotometry in a GeneQuant RNA/DNA Calculator (Pharmacia). The DNA sample was diluted in sterile distilled water to a total volume of 500µl and absorbances at 260 and 280nm were measured. The DNA concentration was calculated according to the formula:

$$\text{DNA concentration} = OD_{260nm} \times Df \times 50,$$

where OD is optical density with OD = 1 = 50µg/ml double stranded DNA, and Df is the dilution factor.

The DNA was used only in a purity higher than 80%, or 2.0 ± 0.2 $OD_{260nm:280nm}$ ratio. Aliquots were diluted to a final concentration of 25µg/ml in sterile distilled water and stored at 4°C.

2.4. Isoenzyme analysis

Leishmania strains were characterized by isoenzyme electrophoretic analysis (IEA), as described by Godfrey and Kilgour (1976) and Harris and Hopkinson (1976), in thin-layer starch-gel. A set of eight enzyme systems that had been found to be polymorphic within the *L. donovani* complex was used: alanine aminotransferase (ALAT, EC 2.6.1.2), aspartate aminotransferase (ASAT, EC 2.6.1.1), glucose-phosphate isomerase (GPI, EC 5.3.1.9), malate dehydrogenase (MDH, EC 1.1.1.37), mannosephosphate isomerase (MPI, EC 5.3.1.8), nucleoside hydrolase using inosine as substrate (NH, EC 3.2.2.1), proline iminopeptidase (PEPD, EC 3.4.11.5) and 6-phosphogluconate dehydrogenase (decarboxylating) (6PGD, EC 1.1.1.44). The general conditions used are presented in Tables VI and VII and were adapted from Evans *et al.* (1984), LeBlancq (1986) and Miles *et al.* (1980a; 1980b),

Tank buffers were prepared the day before and kept at 4°C until use. Starch gels were prepared with 40 ml of diluted buffer (Table VI) and 120% of the amount of starch recommended by the manufacturer, heated until boiling and degassed. The molten gel was then poured onto the plate, spread uniformly and left to cool on a cooling plate. The enzyme extract beads were left to thaw on ice. Cotton thread pieces (number 2) with 0.4 to 0.5mm were soaked in the enzyme extracts and then loaded into the set gel, at 1/3 of the gel plate length. The gel plate was placed in the tank, onto a cooling plate at 4°C, had the centre covered with a glass plate and buffer soaked sponge wicks connected each extremity with the respective buffer filled electrode chamber. The apparatus was held in place by a large glass plate and electrophoresis was run for pre-determined times for each enzyme. For agar development, all gel covers were removed and a frame was placed on the gel, to include the threads and the area where the enzyme bands were expected to be.

Reagents for development (Table VII) were mixed with liquid agar (kept at 56°C), immediately poured onto the gel and left to set. The gel plate was then transferred to a 37°C oven and kept in the dark. For UV development, 1M filter paper, cut to size, was soaked in the mixed developer components and placed over the gel. Upon appearance of bands the gels were photographed using polaroid film, under white or ultra violet (UV) light.

Table VI - Running conditions for thin layer starch-gel electrophoresis of *Leishmania*.

Enzyme	Tank buffer	Gel dilution	Time (hours)	Voltage (V)
ALAT	1	1:9	2	300
ASAT	1	1:9	2	300
GPI	2	1:9	2.5	300
MDH1	3	2:1	4	120
MDH2	4	1:9	4	250
MPI	5	1:4	2.5	300
NHi	6	1:9	2.5	300
6PGD	2	1:9	2.5	300
PEP-D	7	1:7	2.5	300

* - 1.0M $MgCl_2$; ** - 0.1M $MgCl_2$; *** - 0.5$MnCl_2$. Tank buffers, in distilled water: 1 - 0.15M Tris, 7.5mM citric acid, pH 9.0; 2 - 0.1M Tris, 0.1M maleic acid, 1mM magnesium acetate, 1 mM EDTA, pH 7.4; 3 - 0.2M Tris, 1mM EDTA, 47.6mM trisodium citrate, pH 9.5; 4 - 0.25M Tris, 0.09M citric acid, pH 7.0; 5 - 0.1M Tris, 0.01M sodium hydrogen phosphate, pH 7.6; 6 - 0.2M Tris, citric acid to pH 8.9; 7 - 0.15M Tris, 0.1 sodium dihydrogen phosphate, pH 8.2.

Table VII - Staining conditions for thin layer starch-gel electrophoresis of *Leishmania*.

Enzyme	Distil. Water (ml)	Developer Buffer	Activators / inhibitors	Co-enzymes	Linking enzymes	Substrates	Visualization method
ALAT	-	A 6.0ml	-	5.0mg NADH	6.0µl LDH (EC 1.1.27)	12.0mg α-ketoglutaric acid 80.0mg L- alanine	UV, Filter paper
ASAT	-	A 6.0ml	-	5.0mg NADH	6.0µl MDH(EC 1.1.1.37)	12.0mg α-ketoglutaric acid 20.0mg L-aspartic acid	UV, Filter paper
GPI	-	B 13.4ml	* 0.4ml	1.0ml NADP (10.0mg/ml)	10µl G6PDH (1000U/ml)	1.6ml D-fructose-6-phosphate (disodium salt) (10mg/ml)	1.0ml MTT (5mg/ml) ; 1.0ml PMS (5mg/ml) ; 18.4ml Agar (1.2%)
MDH1	3.0	B 8.0ml	-	1.2ml NAD (10.0mg/ml)	-	2.0ml 1M L-malate, pH 7.0	1.0ml MTT; 1.0ml PMS; Agar 16.2ml
MDH2	3.0	B 8.0ml	-	1.2ml NAD (10.0mg/ml)	-	2.0ml 1M L-malate, pH 7.0	1.0ml MTT; 1.0ml PMS; 16.2ml Agar
MPI	-	C 12.0ml	** 2.0ml	0.6ml NADP (10.0mg/ml)	30U GPI (EC 5.3.1.9) type III 20U G6PDH (EC 1.1.1.49)	10mg mannose-6-phosphate	1.0ml MTT; 1.0ml PMS; 16.8ml Agar
NHi	-	D 16.0ml	-	-	50µl xanthine oxidase	20mg inosine	1.0ml MTT; 1.0ml PMS; 18.1ml Agar
6PGD	-	B 12.0ml	** 2.0ml	0.6ml NADP (10.0mg/ml)	-	1.0ml 6-phosphate gluconate (10mg/ml)	1.0ml MTT; 1.0ml PMS; 17.6ml Agar
PEP-D	-	C 14.0ml	*** 0.4ml	-	2.0mg peroxidase (EC1.11.1.7) 2.0mg snake venom	20.0mg L-leucyl L-proline	20mg 3-amino-9-ethyl carbazole in 2ml ethanol; 18.2ml Agar

* - 1.0M $MgCl_2$; ** - 0.1M $MgCl_2$; *** - 0.5$MnCl_2$. Development buffers: A - 0.08M Na_2HPO_4, 0.02M NaH_2PO_4,pH 7.4; B - 0.3M Tris pH 8.0; C - 0.3M Tris pH 7.4; D - 0.3M Tris pH 7.0.

2.5. Random amplification of polymorphic DNA

A set of 28 decanucleotide primers (Table VIII) was assayed for random amplification of polymorphic DNA (RAPD) with the reference strains (*L. donovani* complex and *L. major*) and a selection of *L. infantum* and *L. chagasi* strains. Primers A1-A6, H1-H6 and L1-L6 were obtained from R&D products (Abingdon, UK) and primers D1-D10 were a kind gift from Dr. Douglas Barker (Cambridge University). Seven primers produced a very poor band pattern or none and were discarded (Table VIII). The remaining 21 primers were selected to be used with a large set of strains, of which ten primers (A2, A4, A5, A6, D3, D8, D10, H1, H4, L2), randomly chosen, were used for RAPD analyses. The analysis to study the relationship between *L. infantum* and *L. chagasi*, included three outgroup strains (*L. aethiopica*, *L. major* and *L. tropica*) with a total of 33 strains; 17 *L. infantum*, 11 *L. chagasi* and five *L. donovani*. The analysis of Portuguese *L. infantum* strains included Old World *Leishmania* and *L. chagasi* reference strains and 22 Portuguese *L. infantum* strains. The analysis of *L. donovani* included again Old World *Leishmania* reference strains, the *L. infantum* zymodeme variants and all *L. donovani* strains;46 strains in total.

In order to assess the influence of DNA concentration on RAPDs, 1, 10, 25 and 50 ng of genomic DNA of the *L. donovani* complex reference strains was amplified by RAPD with each of the ten primers used for analysis.

Table VIII - Primers used for RAPD amplification.

NAME	SEQUENCE (5'-3')	NAME	SEQUENCE (5'-3')
A1	tca cga tgc a	D9	AGG TGA CCG T
A2	**GAA ACG GGT G**	**D10**	**GTT GCG ATC C**
A3	AGT CAG CCA C	L1	CGG CCC CTG T
A4	**AAT CGG GCT G**	**L2**	**CGG ACG TCG C**
A5	**CTC ACG TAG G**	L3	CCC GCC ATC T
A6	**CTG ATC GCA G**	L4	GTG GAT GCG A
D1	CAG GCC TTC	L5	AAG AGC CCG T
D2	agg ggt ctt c	L6	aag gat cag a
D3	**TTC CGA ACC C**	**H1**	**CGC GCC CGC T**
D4	GGT CCC TGA C	H2	ttc ccc cgc t
D5	GGG TAA CGC C	H3	cat ccc cct g
D6	CAA TCG CCG T	**H4**	**TGC CGA GCT G**
D7	TCG GCG ATA G	H5	tag gat cag a
D8	**AGC CAG CGA A**	H6	cac atg ctt c

In bold are the primers randomly chosen for the analysis, from among those producing good quality amplification profiles, in lower case the non-working primers.

The RAPD procedure was based on the technique first described by Williams *et al.* (1990) and conditions are described in Annex 1. To ensure homogeneity of the PCR mix throughout the reactions at one time, enough reaction mixture (without DNA) was prepared for the total number of tubes for each RAPD assay (samples and one negative control). Aliquots of the mixture were distributed to the reaction tubes and covered with mineral oil. The DNA was then added through the mineral oil. All strains in each set were amplified at the same time in the same block.

RAPD products were separated by agarose gel electrophoresis: 1.2% agarose in TAE buffer, with 1mg/l EtBr. The RAPD products were run against 1kb molecular weight markers (LKB Pharmacia) in a large gel (25cm x 20cm) using a 42 tooth comb, with 400ml of 1.2% agarose, at 100mV, or in a medium size gel (15cmx12.5cm) using a 20 tooth comb in separate batches, with 100ml of the same gel. The gels were photographed under UV light, initially, in a video gel developing system (Mitsubishi), and later, in a UVP, Inc. (USA) white/UV transilluminator photographic system coupled to programme GRAB-IT, Annotation Grabber 2.51 (Synoptics Ltd, UK). Photographs were printed in a Sony Digital Graphics printer.

2.6. Preparatory PCR amplifications

2.6.1. 3' end of the constitutive major surface protease gene (*MspC3*)

The 3' end of the constitutive major surface protease (*mspC3*), between nucleotides 924 and 2238 of *L. chagasi mspC* (GenBank M80671) was amplified by PCR. Three pairs of primers were designed from an alignment of *L. chagasi mspC*, *mspL* and *mspS1* and *L. donovani* gp63 genes (GenBank, M80671, M80672, M80669 and M60048, respectively). *MspC* specific primers were designed from conserved regions, as judged by comparison with the *L. donovani* sequence, and each pair to amplify 500-650bp. Primers were designed for low repeat content, a GC content of at least 30%, similar melting points (around 60°C), higher GC content in the 3' codon and thymidine residues in the 3' end were avoided. Primers were a Perkin-Elmer product and are shown in Table IX .

Each pair of primers was initally used to amplify three separate regions (A, B, C), but primers C4F and C8R were later used to amplify the full *mspC3* (Fig. 9), which improved the sequence quality and reduced the number of sequences needed for each strain. Each 20µl amplification reaction was optimized for yield and specificity. The optimized PCR conditions are detailed in Annex 1.

Table IX - Primers used to amplify within *mspC* and expected product sizes.

Region	Expected size (bp)	Primer name	Primer sequence (5' - 3')
A	652	C4F	*tgt aaa acg acg gcc agt* cac gtc ggc ttc agt gga
		C6R	*cag gaa aca gct atg acc* aaa aac cct gct gcc aac
B	502	C6F	cca gtc gtc tga tgg tcg
		C7R	ctg ctg gag ctg tcg gag
C	542	C11F	gcg cgg cag tat gga cta
		C8R	tgg acc gga gaa gac gag
mspC3	1351	C4F	(as above)
		C8R	(as above)

R - reverse; F - forward. C4F and C6R had, respectively, the -21M13 and M13Rev primers added (italics).

2.6.2. Gp63 intergenic regions: ITG/CS and ITG/L

Primers for amplification of intergenic regions within the gp63 gene array (Table X, Fig. 10) were designed either to be specific for the intergenic region located upstream of *mspC* (gp63 ITG/CS) or to amplify the intergenic region downstream of any gp63 gene in the array (gp63 ITG), and were designed in the same way as the *mspC3* primers (above). *MspC* specific primer (C9F) was chosen from areas with low or no evidence of variation between sequences for gp63 published for *L. chagasi* and *L. donovani*. Because gp63 ITG/L is multicopy, the *msp* generic primers (C10F and C1R) were designed with added restriction sites, compatible with the multiple cloning site of M13 type plasmids, in order to be able to differentiate different alleles by cloning if necessary. The optimized PCR mixture compositions and temperature cycle conditions for amplification of all intergenic regions described are given in Annex 1.

DNA from *Leishmania* reference strains (Table II) and *L. infantum*, *L. chagasi* and *L. donovani* strains was used in a PCR assay to verify the specificity of the gp63 ITG PCRs. Both undiluted (100-500ng) and 25ng of genomic DNA were used.

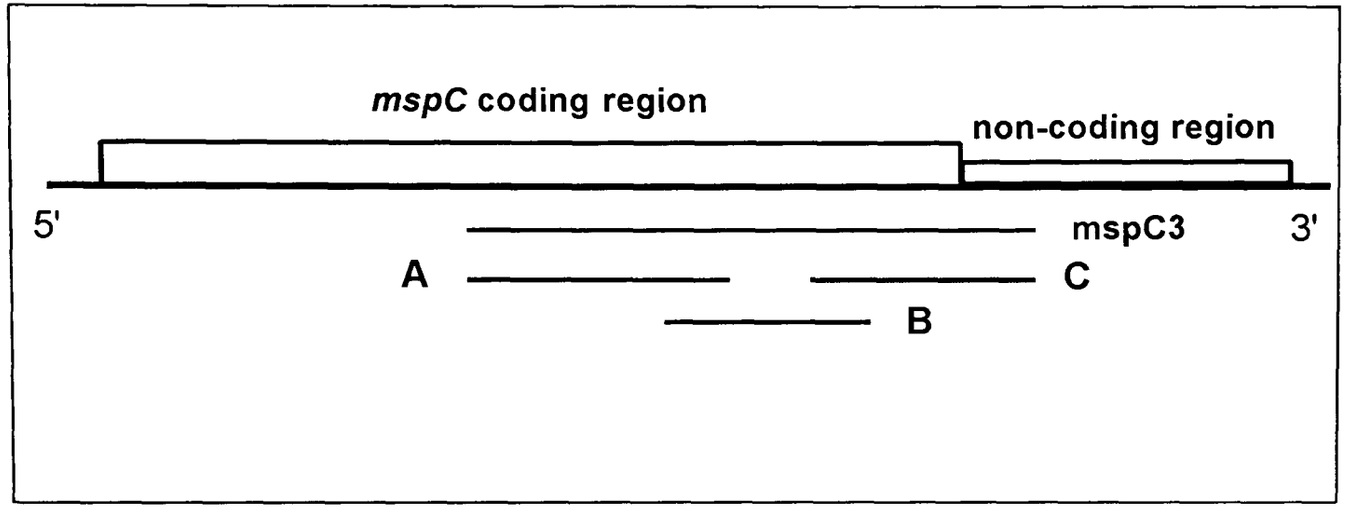


Figure 9 - Amplified regions within the sequence of the *L. chagasi mspC* (GenBank M80671).

Table X - Primers designed for PCR amplification of gp63 ITGs.

Region	Expected size (kb)	Name	Sequence (5' - 3')	Site (nt) of 3' end in GenBank M80671	Insert
gp63 ITG/L	1.7 + 2.0 (/ 4.5)	C10F	**ggg aag ctt** acg tac agc gtg cag gtg	1697	*Hind*III
		+			
		C1R	**ggg ccc ggg** cga cag cag cga tga ctg	210	*Xma*I
gp63 ITG/CS	~1.6				
		+			
		C9F	ggc tcc cga cgt gag tta	2055	-

*Hind*III, *Xma*I - sequences for restriction with the named enzymes (bold). R - reverse; F - forward.

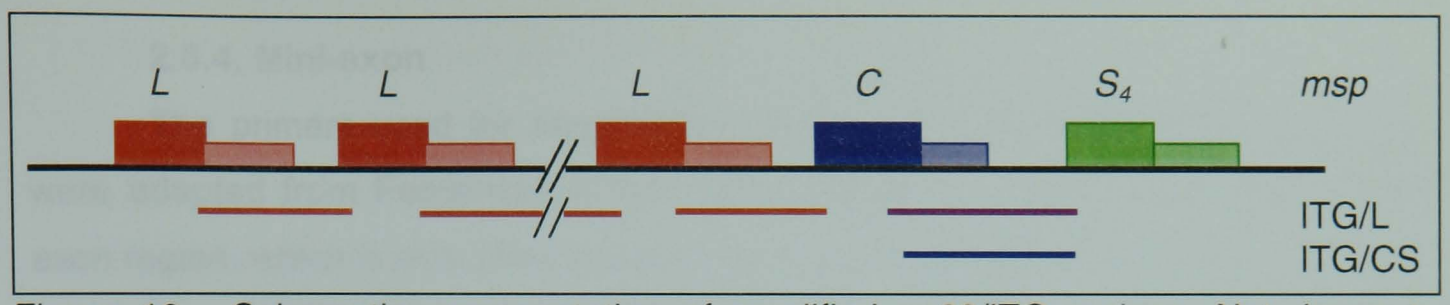


Figure 10 - Schematic representation of amplified gp63/ITG regions. Not drawn to scale. Large and small rectangles are, respectively, coding and non-coding regions. // represent and unknown number of of *mspL* unit repeats. The purple line represents an ITG/CS product generated by the ITG primers used to amplify ITG/L, whilst the red lines are the required ITG/L product. The blue line is the ITG/CS product generated from the specific ITG/CS PCR.

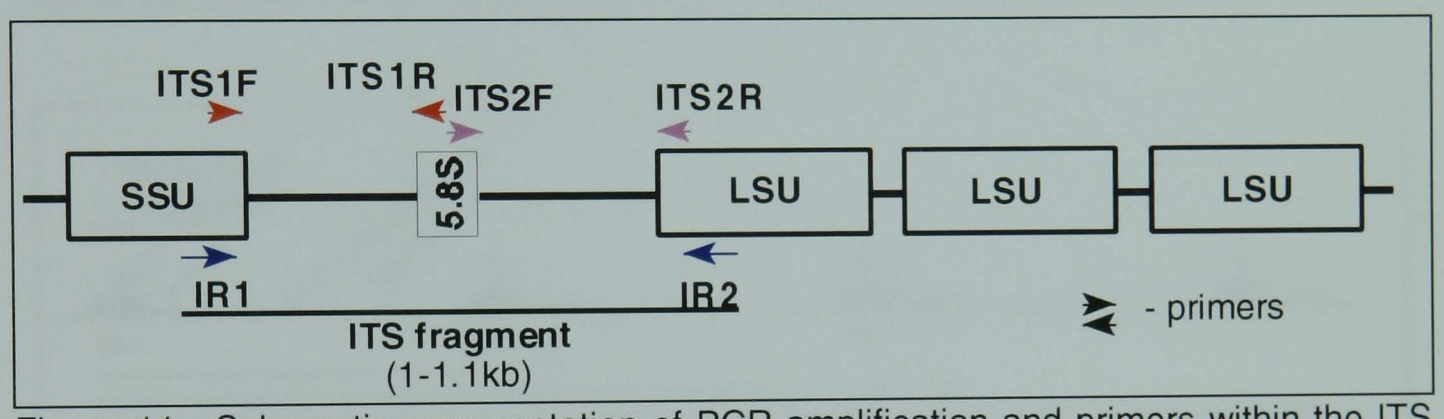


Figure 11 - Schematic representation of PCR amplification and primers within the ITS region (not drawn to scale). SSU - small sub-unit; LSU - large sub-unit; ITS - internal transcribed spacer.

2.6.3. Internal transcribed spacer (ITS)

For amplification of the ribosomal internal transcribed spacer (ITS), the same primers were used as described by Cupolillo *et al.* (1995):

IR1 - 5' GCT GTA GGT GAA CCT GCA GCA GCT GGA TCA TT 3' Forward

IR2 - 5' GCG GGT AGT CCT GCC AAA CAC TCA GGT CTG 3' Reverse

The initial PCR conditions (Annex 1, ITSa) were derived from Stothard *et al.* (1996) but later modified (ITSb): $MgCl_2$ was reduced to 3mM and primers to 30nM. The expected PCR products for OW *Leishmania* were between 1.1 and 1.0kb.

Four internal primers to the ITS PCR products (Fig. 11) were designed, from an alignment with all trypanosomatid ribosomal DNA sequences available, in order to sequence the variable intergenic regions in the *Leishmania* ITS:

ITS1F	5' GCA GCT GGA TCA TTT TCC 3'	Forward
ITS1R	5' AGC TTC TCC CAT GCG CCG 3'	Reverse
ITS2F	5' GTC ATC CCC GTG CAT GCC 3'	Forward
ITS2R	5' AAC ACT CAG GTC TGT AAA C 3'	Reverse

2.6.4. Mini-exon

The primers used for amplification of the mini-exon repeat unit (Figure 12), were adapted from Fernandes *et al.* (1994). Primers are located in the conserved exon region, which is only 39nt, and thus were partially complementary in their 5'end.

ME1 - 5' CAA TAT AGT ACA GAA ACT G 3' Reverse

ME2 - 5' TTC TGT ACT TTA TTG GTA 3' Forward

PCR amplification of the mini-exon repeat unit (ME/unit) was done according to Ramos *et al.* (1996) (Annex 1) with expected products for Old World *Leishmania* between 400 and 460bp. In order to improve the PCR product and reduce the amount of primer used, protocols were modified to include only 30 cycles, and to use 1/4 of the primer concentration with twice the amount of DNA, respectively .

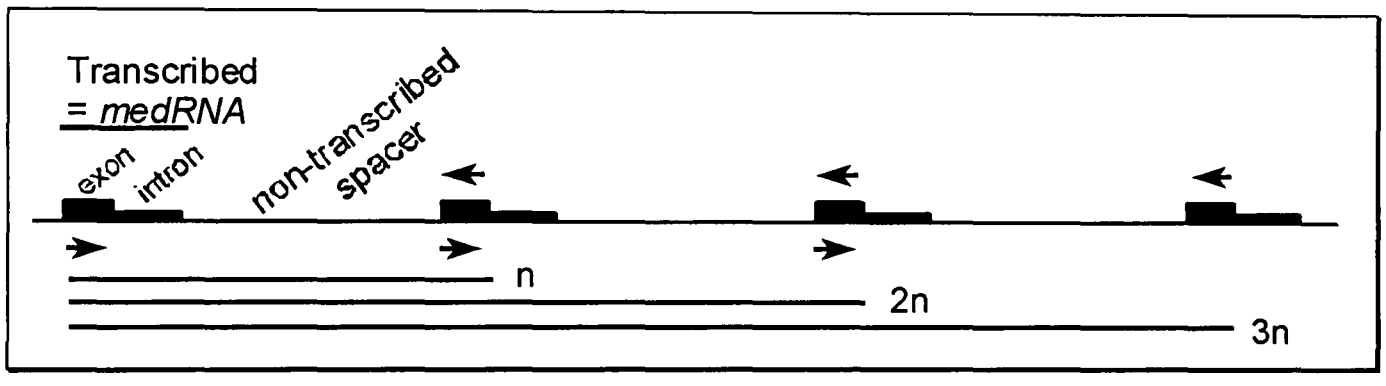


Figure 12 - Schematic representation of the mini-exon PCR amplification. The primers are indicated by arrows. Because primers are partially complementary, multiple unit repeats (2n, 3n) were also expected in addition to the single unit amplification (n).

2.6.5. Preparation of PCR products

The gp63 ITG/L band (1.6 kb) was separated and purified from the other ITG PCR products, the mini-exon unit fragment was purified from multiple amplification fragments and all other PCR products were also purified to eliminate spurious non-specific background and primers. After separation in 0.8% agarose gels the required PCR fragments were excised and purified using a QIAEX II™ gel purification Kit (Qiagen), according to the manufacturer's instructions. The purified DNA was kept in 10mM Tris-HCl buffer, pH 8.5.

Different concentrations of DNA template were sequenced initially to assess the best sequencing concentration. The following sequence reactions used 50 to 100ng DNA as calculated, or extrapolated, by comparing with a working reference in a 1% agarose gel. The template DNA concentrations for RFLP analysis were adjusted to roughly 0.1μg/μl amongst all strains by visual estimation in an agarose gel.

2.7. Restriction fragment length polymorphisms (RFLPs)

The ITS, mini-exon and both gp63 ITG PCR products were analysed through enzyme restriction. Restriction enzymes (Table XI), with a minimum of two restriction sites in each target sequence, were chosen either by trial and error with six Old World WHO reference strains (ITS and gp63 ITGs) and/or by screening published sequences (gp63 ITGs and mini-exon). Restriction reactions were of 10μl, with approximately 0.5μg of DNA and 5U of restriction enzyme in the recommended buffer and were incubated overnight at 37°C (or at 65°C for *Taq*I, 60°C for *Bst*UI). Restriction products were separated at 250V for 2h in 6% acrylamide, 3.5% cross linking gels, for good resolution of smaller DNA fragments: 40ml 10.5% volume/volume (v/v) 2% liquid bisacrylamide, 20% v/v 30% liquid acrylamide (National Diagnostics) in TBE (Tris-borate-EDTA buffer: 89mM Tris-borate, 2mM EDTA), filtered and degassed, and then were added, per plate, 70μl TEMED (NNN'N' tetramethylethylenediamine), 260μl 10% weight/volume (w/v) ammonium persulphate. The gels were silver stained (Stothard *et al.*, 1998). Briefly, the gels were twice fixed for 3min in 10% v/v ethanol, 0.5% acetic acid; stained for 10min in 0.1% w/v silver nitrate ($AgNO_3$); washed twice in distilled water; developed in 1.5% w/v sodium hydroxide (NaOH), 0.01% w/v sodium borohydrate ($NaBH_4$), 0.015% v/v formaldehyde, until bands were seen and neutralized for 10min in 0.75% w/v sodium carbonate (Na_2CO_3). All incubations were done with 300ml of each solution with gentle agitation. The gels were stored in plastic bags and photographed under white light with the UVP system (see 2.5).

Table XI - Restriction enzymes used for analysis of PCR amplification products.

Enzyme	Recognition site	ITS	Gp63 ITGs	Mini-exon
*Aci*I	C.CGC	-	-	A
*Alu*I	AG.CT	A	A	-
*Bst*UI	CG.CG	A	A	A
*Cfo*I	GCGC	A	A	A
*Hae*III	GG.CC	A	A	A
*Mse*I	T.TAA	A	-	-
*Msp*I	C.CGG	-	A	A
*Sau*3AI	.GATC	-	-	-
*Scr*FI	CC.NGG	-	-	A
*Rsa*I	GT.AC	-	A	-
*Taq*I	T.CGA	A	A	-
*Eco*RI	G.AATTC	A	-	-
*Sph*I	GCATG.C	A	-	-

A - enzyme applied. The dot represents the cleavage site for each enzyme.

2.8. DNA sequencing

Purified *mspC* PCR products, obtained from selected strains, were sequenced. A partial gp63 ITG/L sequence was also obtained for *L. chagasi* PP75.

Direct sequencing, of either full *mspC3* or partial PCR products (A, B, C), was done with ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction kits (Perkin-Elmer, UK) or AmpliTaq® DNA Polymerase, FS Thermo-sequenase Dye Terminator Cycle Sequencing Premix kits (Amersham, Life Science) according to the manufacturers' instructions, using the PCR primers or internal primers. The cycling protocol was, 25 cycles at 96°C for 30s, 50°C for 15s, 60°C for 240s. The sequencing products were purified by ethanol precipitation with 3M sodium acetate pH 4.6, according to the sequencing kit manufacturer's instructions, and separated in an ABI Prism™ 377 DNA Sequencer (Perkin-Elmer, UK). Consensus sequences were obtained from forward and reverse reactions aligned in ABI Prism Sequence Navigator™ Version 1.0.1 (Perkin-Elmer, UK) using Clustal V (Higgins *et al.* 1992).

2.9. Phylogenetic analysis

2.9.1. Analysis of band patterns

Isoenzyme bands were scored for analysis as presence / absence data, but also coded as nucleotide residues or amino acids in, respectively, DNA or protein sequence data. Within the *L. donovani* complex, only A and T letters were used to minimize the influence of base composition on the results. Algorithms which did not categorize bases or aminoacids were used when possible.

Electrophoretic bands from isoenzyme analysis, RAPD or RFLP were scored for presence (1) or absence (0), except RFLP fragments larger than 900 bp for which resolution was poorer. Bands that were too faint or judged inconsistent and strains with poor RAPD amplification profiles were ignored. Data matrices of each individual RAPD primer or each PCR-RFLP were analysed separately, or as pooled data, with packages SYN-TAX-pc 5.0 (Podani, 1993) or PHYLIP (Felsenstein, 1993).

Dendrograms were built from a Jaccard distance matrix {$\sqrt{[1-a/(a+b+c)]}$, in which a is a double presence and b and c are presence in one taxa and absence in the other} using single linkage and unweighted pair group method with arithmetic averages (UPGMA, SYN-TAX), neighbour-joining and Fitch-Margoliash (PHYLIP) clustering algorithms. In some cases minimum spanning trees were superimposed on principal coordinates distributions. Cophenetic correlation was used to verify the fit of tree derived distances (SYN-TAX) to observed pairwise distance matrix. The cophenetic correlation coefficient ranges from 0-1, in which 1 is the perfect fit.

PHYLIP analyses - neighbour-joining and UPGMA (NEIGHBOR), Fitch-Margoliash (FITCH), Wagner and Camin-Sokal parsimony (MIX), polymorphism and Dollo (DOLLOP) parsimonies - were performed using taxa jumbling and global rearrangements. Parsimony bootstrap analyses were done for 100 replications to estimate robustness of lineage divergence. Consensus trees were obtained by majority rule consensus (CONSENSE).

2.9.2. Analysis of *mspC3* DNA sequences

The sequences of *L. donovani* complex *mspC3*, and the putative homologues *L. major gp63-6* and *L. mexicana gp63-C1*, from GenBank (respectively AF039721 and X64394), in a total of 1083, bases were aligned by Clustal V (Higgins *et al.* 1992) and were used to infer phylogenies using programmes within the PHYLIP package (Felsenstein, 1993). Dendrograms were produced from sequence alignment data using maximum parsimony (DNAPARS) and maximum likelihood methods (DNAML), and from a distance matrix (corrected nucleotide divergence using Kimura-2-parameter model; DNADIST) using neighbour-joining and Fitch-Margoliash least squares methods. Bootstrap analyses were done, except for the maximum likelihood method, based upon 1000 or 100 replicate data sets.

3. Isoenzyme Characterization

3.1. Introduction

3.1.1. Isoenzyme analysis

The ultimate phenotypic characters are proteins, responsible for morphology and metabolism, and which are direct products of genes although they may be subject to post translational modifications. Proteins thus reflect DNA sequence and composition, even though phenotype may vary with environmental factors and with cell differentiation.

Protein electrophoresis allows distinction between different proteins (allelic products) mainly by electrical charge but also by size or conformation (Avise, 1994). Not all proteins are amenable to analysis; most membrane proteins cannot easily be analysed by electrophoresis because of hydrophobic regions. Proteins can be visualised with a general protein detection strain, but only those that are abundant will be easily identified (May, 1992). Proteins present in small amounts and those which cannot be identified by size alone require specific identification techniques. Monoclonal antibodies designed to conserved regions can be used to identify the desired protein (eg. by western-blotting). Most enzymes, most notably metabolic enzymes, are soluble and catalyze specific reactions which can be used for direct identification on an electrophoretic gel.

Specific enzyme staining can be achieved through production of coloured or fluorescent products either of the reaction itself for by coupling to a dye reaction (May, 1992). Each enzyme reaction will produce a banding pattern which is characteristic of the enzyme and organism and is dependent on the quaternary structure of the enzyme and number of alleles and loci. Sub-units in the quaternary structure of the enzyme, multiple alleles and multiple loci give rise to different motilities increasing and complicating the number of bands seen in the heterozygous profile (May, 1992). Analysis can be further complicated by generation of artifacts, either by conformational changes in the enzyme due to electrophoretic or extraction conditions or by expression of an alternate enzyme with similar reaction requirements, or by contamination (May, 1992).

The study of electrophoretic enzyme banding patterns, known as allozyme or, more broadly, isoenzyme analysis (IEA), has many applications, such as the study of organ metabolism, allele and gene composition and organization, quaternary structure of enzymes, gene linkage and determination of allele frequencies in populations (Avise, 1994; May, 1992). The application most relevant to this project is in typing of organisms (May, 1992). Some diversity will remain undetected, however,

because only variation producing charge differences are evident. Despite this limitation, IEA is widely used because it was the first method to produce reliable taxonomic data for *Leishmania* and thus it became the reference method for identification of *Leishmania*, as well as for many other unicellular organisms.

IEA resolution can be improved for some enzymes by isoelectric focusing (IEF) (Piarroux *et al.*, 1995) on a gel with a pH gradient (May, 1992). IEF allows better discrimination by separating proteins on the basis of their isoelectric point, but it has some limitations, because some enzymes lose activity when subject to IEF conditions. However, IEF data can complement standard isoenzyme characterization and genotypic variance data (Piarroux *et al.*, 1995).

Media other than starch are also used for allozyme electrophoresis, such as cellulose acetate and polyacrylamide. Cellulose acetate is faster and requires less sample and staining reagents (May, 1992), but fewer samples can be loaded. Polyacrylamide gels are adequate for IEF and, by addition of SDS (a detergent), proteins and their sub-units can be separated according to size. The highest throughput method is thick starch gel electrophoresis, as gels can be sliced and thus used for detection of several enzymes simultaneously.

3.1.2. Isoenzyme typing of *Leishmania*

Many enzymes, mainly from the glycolytic pathway, are available for typing *Leishmania* and, upon development of the allozyme electrophoresis method (Godfrey and Kilgour, 1976; Harris and Hopkinson, 1976), different enzymes and methods were introduced. The LSHTM reference laboratory, uses thin layer starch gel electrophoresis with 12 enzymes, of which 8 were found to discriminate between strains of the *L. donovani* complex (underlined): alanine aminotransferase (ALAT; E.C.2.6.1.2), aspartate aminotransferase (ASAT; E.C.2.6.1.1), esterase (ES; E.C.3.1.1.1), glucosephosphate isomerase (GPI; E.C.5.3.1.9), malate dehydrogenase (MDH; E.C.1.1.1.37), mannosephosphate isomerase (MPI; E.C.5.3.1.8), nucleoside hydrolase (NH; E.C.3.2.2.2), 6-phosphogluconate dehydrogenase (6PGD; E.C.1.1.1.44), proline iminopeptidase (PEP-D; E.C.3.4.11.5), phosphoglucomutase (PGM; E.C.2.7.5.1), pyruvate kinase (PK; E.C.2.7.1.40) and superoxide dismutase (SOD; E.C.1.15.1.1). Zymodemes, or enzyme types, from the LSHTM are coded with serial numbers and the prefix LON. Under this system all *L. infantum* were typed as LON49, *L. donovani* ss defined as LON 41 with 10 other zymodemes described as *L. donovani* sl. The Montpellier reference laboratory uses a system of 15 enzymes, which are suitable for described *Leishmania* species and for construction of phylogenies (Rioux *et al.*, 1990): MDH, malic enzyme (ME; E.C.1.1.1.40), isocitrate

dehydrogenase (ICD; E.C.1.1.1.42), PGD, glucose-6-phosphate dehydrogenase (G6PD; E.C.1.1.1.49), glutamate dehydrogenase (GLUD; E.C.1.4.1.3), diaphorase NADH (DIA; E.C.1.6.2.2), purine nucleoside phosphorylase 1 (NP1; E.C.2.4.2.1), purine nucleoside phosphorylase 2 (NP2; E.C.2.4.2), glutamate-oxaloacetate transaminase (GOT1, GOT2; E.C.2.6.1.1), PGM, fumarate hydratase (FH; E.C.4.2.1.2), MPI and GPI. Montpellier uses a typing code with the prefix MON. MON 1 is for the most common *L. infantum* (LON 49), which comprises more than 17 other zymodemes (Rioux *et al.*, 1990) further increased with the discovery of new zymodemes in HIV patients. *L. donovani* reference strain is zymodeme MON 2 and 12 more zymodemes have been described.

3.1.3. Technical procedure

Enzymes are released by rupture of the cell membranes either by mechanical action or freeze-thawing, into a stabilizing buffer. Extracts are stored in liquid nitrogen in beads of 10µl, which, when needed, are each placed in a well of a multiwell plate to thaw and kept on ice during use.

Preparation of a thin layer starch gel is achieved by boiling starch in diluted tank buffer, pouring and evenly spreading on glass plates. Gels must be free of any gas bubbles and are placed on a cooling plate. Small cotton threads are soaked in the samples and loaded onto the gel on a straight line. Plates are covered and connected to the tank buffer by soaked sponge wicks. Adequate times, voltages and buffers for electrophoresis have to be selected empirically. Staining reagents are mixed and poured onto the gel as a solution for colour staining or soaked into filter paper for fluorescent staining. Staining gels are poured into frames placed on the gel and the filter paper is placed directly onto the gel. Gels are incubated until bands are optimally stained for photography.

3.1.4. Future of isoenzyme analysis

Isoenzyme analysis still proves useful in many areas of biology and medicine, but there are many disadvantages, especially for *Leishmania*: only limited genetic diversity can be identified, only a few loci are usually examined, only strains isolated from the host and free of contaminants can be studied, and a large number of parasites are required for analysis. Until now, IEA has been cheaper and easier to perform than genotyping, and was used to gather the most comprehensive genetic data on *Leishmania*. The entire *Leishmania* genome is being sequenced, however, and genotyping methods are being developed that are faster, cheaper and can be species specific such that isolation and culture of parasites are not necessary. Most

PCR based methods are also sensitive enough to require only small samples. By genotyping it may also be possible to identify parasites from host biopsies without the lengthy and hazardous process of culture.

Despite the recent focus on genomics, protein analysis expanded to the entire protein complement of the genome (proteomics). The sequencing of genomes allows faster identification of proteins and thus of their relation with biological function.

3.2. Results

3.2.1. Enzyme profiles for each strain

Each strain was analyzed by thin layer starch gel electrophoresis and stained for the enzymes: ALAT, ASAT, GPI, MDH, MPI, NH, PEPD, 6PGD. According to Le Blancq (1986), these were the polymorphic enzymes for the *L. donovani* complex among the 12 which have been used in the LSHTM to characterize *Leishmania* strains. The profiles obtained for each strain and the codes from previous characterizations are listed for the outgroup strains in table XII, *L. chagasi* strains in table XIII, *L. infantum* strains in table XIV and *L. donovani* strains in table XV. Graphical depictions of the profiles for each enzyme are shown in Figure 13.

In the present work, some profiles did not match those described by Le Blancq (1986). The LON 45 ALAT double band profile which had been described by Le Blancq (1986) was not consistent because a similar pattern was also observed for other strains. Each ALAT profile presented only one consistent single band, but two other alleles were detected in strains C1 (WR341), D23 (Sukkar 2) and D26 (Wangjie 1) (Tables XIII and XV, and Figure 13). Since zymodeme LON 45 could also be differentiated from LON zymodemes by other enzymes, its designation was kept in the absence of the ALAT characteristic pattern, but in the form of LON 45*.

Two single band 6PGD profiles were identified in the *L. donovani* complex, and according to Le Blancq (1986) an extra, faint, band of the same size as the slower 'allele' should have been present in all strains with the faster band pattern, except LON 52. In this work only a few of the expected strains had the extra band, therefore, and because the phenotype was not consistent with conventional heterozygosity, only the stronger bands were scored. It became thus impossible to differentiate between zymodemes LON 42 and LON 52, which are here referred to as LON 42/52.

It was very difficult to discriminate between MDH patterns, especially between the fainter bands but also between the sizes of the slower stronger bands. For simplicity, only the stronger bands were coded and the profiles corresponding to relative motilities 100 and 104 in the MON system were not distinguished.

Table XII - Enzyme profiles of the outgroup strains.

Code	Name	Enzymes*								Zymodeme		
		ALAT	ASAT	GPI	MDH	MPI	NH	PEPD	6PGD	LON	MON	ILM
L. aethiopica	MHOM/ET/1970/L96	E	E	D	D	E	F	E	E	27	14?	18
L. major	MHOM/SU/1973/5-ASKH	D	C	C	-	D	E	D	D	1	4	17
L. tropica	MHOM/SU/1974/K27	B	-	B	A	-	-	F	-	-	60	-

* distinct isoenzyme profiles are designated A, B, C, etc.

Table XIII - Enzyme profiles of *L. chagasi* strains.

Code	Name	Enzymes*								Zymodeme		
		ALAT	ASAT	GPI	MDH	MPI	NH	PEPD	6PGD	LON	MON	ILM
C	MHOM/BR/1974/PP75†	A	C	A	A	A	B	A	B	49	1	9
C3	MCAN/BR/1984/CO910	-	C	-	-	A	B	-	B	-	-	-
C8	MCAN/BR/1989/DOG118	A	C	A	A	A	B	A	B	-	-	9
C7	MCAN/BR/1989/DOG136	A	C	A	A	A	B	A	B	-	-	9
C13	MCER/BR/1981/M6445	A	C	A	A	A	B	A	B	-	-	9
C9	MCER/BR/1983/M7633	-	C	-	-	A	B	-	B	-	-	-
C12	MCER/BR/1989/M12084	-	C	-	-	A	B	-	B	-	-	-
C11	MCER/BR/1989/M1201985	-	C	-	-	A	B	-	B	-	-	-
C5	MHOM/BR/1984/M8270	-	C	-	-	A	B	-	B	-	-	-
C4	MHOM/BR/1985/M9702	-	C	-	-	A	B	-	B	-	-	-
C2	MHOM/PA/1978/WR285	A	C	A	A	A	B	A	B	-	-	9
C1	MHOM/PA/1980/WR341	C	A	A	A	A	B	A	B	(204)	-	14

Not all strains were fully typed. In parenthesis is a LON zymodeme assigned in this project, based on partial characterization. † WHO reference strain. * distinct isoenzyme profiles are designated A, B, C, etc.

Table XIV - Enzyme profiles of *L. infantum* strains.

Code	Name	Enzymes*								Zymodeme		
		ALAT	ASAT	GPI	MDH	MPI	NH	PEPD	6PGD	LON	MON	ILM
I	MHOM/TU/1980/IPT-1*†	A	C	A	A	A	B	A	B	49	1	9
I23	IARI/PT/1989/IMT 169	-	C	-	A	A	B	A	B	-	1	-
I24	IARI/PT/1989/IMT 170	-	C	-	-	A	B	-	B	-	1	-
I25	IARI/PT/1989/IMT 171	A	C	A	A	A	A	A	B	(201)	24	11
I26	IARI/PT/1989/IMT 172	A	C	A	A	A	A	A	B	(201)	24	11
I3	MCAN/FR/1982/Pharoah	A	C	A	A	A	B	A	B	49	-	9
I1	MCAN/PT/1981/L82	A	A	A	A	A	B	A	B	49**	-	8
I10	MCAN/PT/1981/IMT 89	A	C	A	A	A	B	A	B	49	-	9
I11	MCAN/PT/1982/IMT 124	A	C	A	A	A	B	A	B	-	-	9
I8	MCAN/PT/1987/IMT 150	-	-	-	-	-	-	-	-	-	1	-
I9	MCAN/PT/1987/IMT 152	A	C	A	A	A	B	A	B	-	1	9
I2	MCAN/PT/1988/Rebelo 2	A	C	A	A	A	B	A	B	-	-	9
I20	MCAN/PT/1989/IMT 160	A	C	-	A	A	B	A	B	-	1	-
I22	MCAN/PT/1989/IMT 161	-	C	-	A	A	B	A	B	-	1	-
I21	MCAN/PT/1989/IMT 162	-	C	-	A	A	B	A	B	-	1	-
I27	MCAN/PT/1991/IMT 177	-	C	-	-	A	B	-	B	-	1	-
I29	MCAN/PT/1993/IMT 191	-	C	-	-	A	B	-	B	-	1	-
I13	MCAN/PT/1993/IMT 193	-	C	-	-	A	B	-	B	-	1	-
I30	MCAN/PT/1994/IMT 195	A	C	A	A	A	B	A	B	-	1	9
I18	MCAN/PT/1994/IMT 204	-	C	-	-	A	B	-	B	-	1	-
I19	MCAN/PT/1994/IMT 205	-	C	-	-	A	B	-	B	-	1	-
I16	MHOM/CN/1980/Strain A	A	C	A	A	A	B	A	B	-	34	9
I17	MHOM/CY/1963/L53	A	B	A	A	A	A	A	B	(202)	-	12
I4	MHOM/ES/1987/Lombardi	A	C	A	A	A	A	A	B	-	-	11
I5	MHOM/FR/1978/LEM75	-	C	A	A	A	B	A	B	-	1	-
I15	MHOM/IT/1981/Alessandro	A	C	A	A	A	B	A	B	-	1	9
I31	MHOM/MT/1985/Buck	A	C	A	A	A	A	A	B	49**	78	11
I6	MHOM/PT/1982/IMT 104	A	C	A	A	A	B	A	B	-	1	9
I12	MVUL/PT/1983/IMT 108	-	-	-	-	-	-	-	-	49	-	-
I7	MVUL/PT/1984/IMPT 128	-	-	-	-	-	-	-	-	-	1	-

Not all strains were fully typed. In parenthesis are LON zymodemes assigned in this project, based on partial characterization. † WHO reference strain. * distinct isoenzyme profiles are designated A, B, C, etc. ** is a zymodeme previously reported by Le Blancq (1986) but found here to be different.

Table XV - Enzyme profiles of *L. donovani* strains.

Code	Name	Enzymes*								Zymodeme		
		ALAT	ASAT	GPI	MDH	MPI	NH	PEPD	6PGD	LON	MON	ILM
D	MHOM/IN/1980/DD8[*†]	A	A	A	A	A	A	A	A	41	2	1
D28	MHOM/ET/1972/Gebre 1[†a]	A	B	A	B	A	A	A	B	50	82	6
D21	IMAR/KE/1962/LRC-L57	A	A	A	A	A	A	A	B	44	37	3
D35	ISER/PT/1993/IMT 188	A	A	A	B	A	A	A	B	-	-	5
D22	MARV/SD/1962/LRC-L64	A	C	A	B	A	A	A	B	48	-	7
D24	MCAN/IT/1976/Dora	A	B	A	B	A	A	A	B	50	-	6
D25	MCAN/KE/0000/D2	A	A	A	A	A	B	A	B	45	-	8
D23	MCAN/IQ/1981/Sukkar 2	B	D	C	C	C/E	D	C	C	43**	-	13
D26	MHOM/CN/0000/Wangjie-1	B	A	B	A	A	C	A	B	42**	35	16
D12	MHOM/ET/0000/Ayele 5	A	A	B	A	B	A	B	B	52	-	10
D27	MHOM/ET/0000/Ayele 8	A	A	A	A	A	A	A	B	56	-	3
D13	MHOM/ET/0000/Hussen	A	A	B	A	B	A	B	B	42	-	10
D1	MHOM/ET/1967/HU3 (LV9)[†]	A	A	A	B	A	A	A	B	46	18	5
D15	MHOM/ET/1984/Addis 142	A	A	A	A	A	A	A	B	-	-	3
D29	MHOM/ET/1984/Addis 164	A	A	B	A	B	A	B	B	-	83	10
D14	MHOM/ET/1982/Bekele	A	A	B	A	C	A	B	B	42	-	10
D6	MHOM/IN/1977/Chowd-X	A	A	A	A	A	A	A	A	-	-	1
D7	MHOM/IN/1979/STL1-79	A	A	A	A	A	A	A	A	-	-	1
D8	MHOM/IN/1982/Nandi 1	A	A	A	A	A	A	A	A	41	-	1
D4	MHOM/IN/1982/Patna 1	A	A	A	A	A	A	A	A	41	-	1
D2	MHOM/KE/1967/MRC(L)3	A	A	A	A	A	A	A	B	-	-	3
D3	MHOM/KE/1973/MRC74	A	A	A	A	A	A	A	A	51	-	1
D31	MHOM/KE/1975/Mutinga H9	A	A	A	A	A	A	A	B	56	32	3
D16	MHOM/KE/1980/Ndandu 4A	A	A	A	A	A	A	A	B	44	-	3
D30	MHOM/KE/0000/Neal-R1	A	A	A	A	A	A	A	B	56	-	3
D17	MHOM/LB/1984/Salti 4	A	A	A	B	A	A	A	B	-	-	5
D34	MHOM/PT/1992/IMT 180	A	A	A	B	A	A	A	B	-	18	5
D32	MHOM/SA/1981/Jeddah KA	A	A	B	A	B	A	B	B	42	31	10
D9	MHOM/SA/1987/VL23	A	A	B	A	B	A	B	B	-	-	10
D10	MHOM/SA/1987/VL29	A	A	B	A	B	A	B	B	-	-	10
D11	MHOM/SA/1987/VL6	A	A	B	A	B	A	B	B	-	-	10
D18	MHOM/SD/0000/Khartoum	A	A	A	B	A	A	A	B	46	18	5
D33	MHOM/SD/1982/Gilani	A	C	A	B	A	A	A	B	48	30	7
D19	MHOM/SD/1985/A22	A	C	A	B	A	A	A	B	-	-	7
D20	MHOM/SD/1987/UGX-marrow	A	A	B	A	B	A	B	B	-	-	10
D5	MMER/IR/1996/MESH-17	-	-	-	-	-	-	-	-	50	-	6

† WHO reference strain. [a] *L. archibaldi*. * distinct isoenzyme profiles are designated A, B, C, etc. ** is the zymodeme previously reported by Le Blancq (1986) but found here to be different.

3.2.2. Zymodeme assignments

The zymodeme numbers assigned to each strain are listed for the outgroup strains in table XII, *L. chagasi* strains in table XIII, *L. infantum* strains in table XIV and *L. donovani* strains in table XV, with the graphical depiction of the profiles of each enzyme and each zymodeme in Figure 13. A new coding system (ILM) was implemented because of the incompatibilities between the LON coding system and the procedure used here. More specifically, ILM uses partial typing (eight enzymes) and a simpler ALAT, 6PGD and MDH profile coding. Furthermore, some strains were found here to have different profiles from those which had been previously described (Le Blancq, 1986). Fifteen distinct profiles were described (Fig. 13; ILM 1, 3, 5-14, 16-18), but three more numbers (ILM 2, ILM 4 and ILM 15) were also assigned, which could correspond to distinct LON zymodemes (Fig. 13) if MDH had been coded as for the LON system.

Five novel zymodeme profiles in relation to previously described LON zymodemes (Le Blancq, 1986) are described here. Strain Sukkar 2 (D23) was assigned zymodeme ILM 13. This strain was heterozygous for 6PGD, had an ALAT profile similar to strain Wangjie 1 (D26) and had a GPI profile similar to *L. major*, but all other assayed enzymes were different from any strain tested in the *L. donovani* complex and outgroups. Strain Wangjie 1 (D26) was typed as ILM 16, which was similar to ILM 10, except for an ALAT profile similar to ILM 13 (see above) and a novel distinctive NH profile. Strain WR341 was assigned zymodeme ILM 14, which was very similar to ILM 8, although with a novel ALAT profile. Some *L. infantum* strains have two different novel combinations of previously known alleles which were coded as ILM 11 and ILM 12. ILM 12 had the particularity of being, as ILM 6, heterozygous for ASAT. The profiles observed for all zymodemes are shown in Figure 13. The new zymodemes (ILM 11-16) were also tentatively assigned a LON number, starting from 201 and are described here with the respective profiles.

Zymodemes MON 24, 34, 35 and 83 were typed by the LON system for the first time, although using only eight enzyme systems. The profiles observed here were equivalent to those observed for the same enzymes by Rioux *et al.* (1990), except for MDH 100 and 104, which could not be discriminated from each other, and ASAT which intermediate MON code (113) was seen and coded here as a heterozygous profile. MON 24 and MON 35 strains were typed as novel zymodemes ILM 11 and ILM 16, respectively, whilst MON 34 and MON 83 strains were indistinguishable from previously described zymodemes ILM 9 (LON 49) and ILM 10 (LON 42/52), respectively.

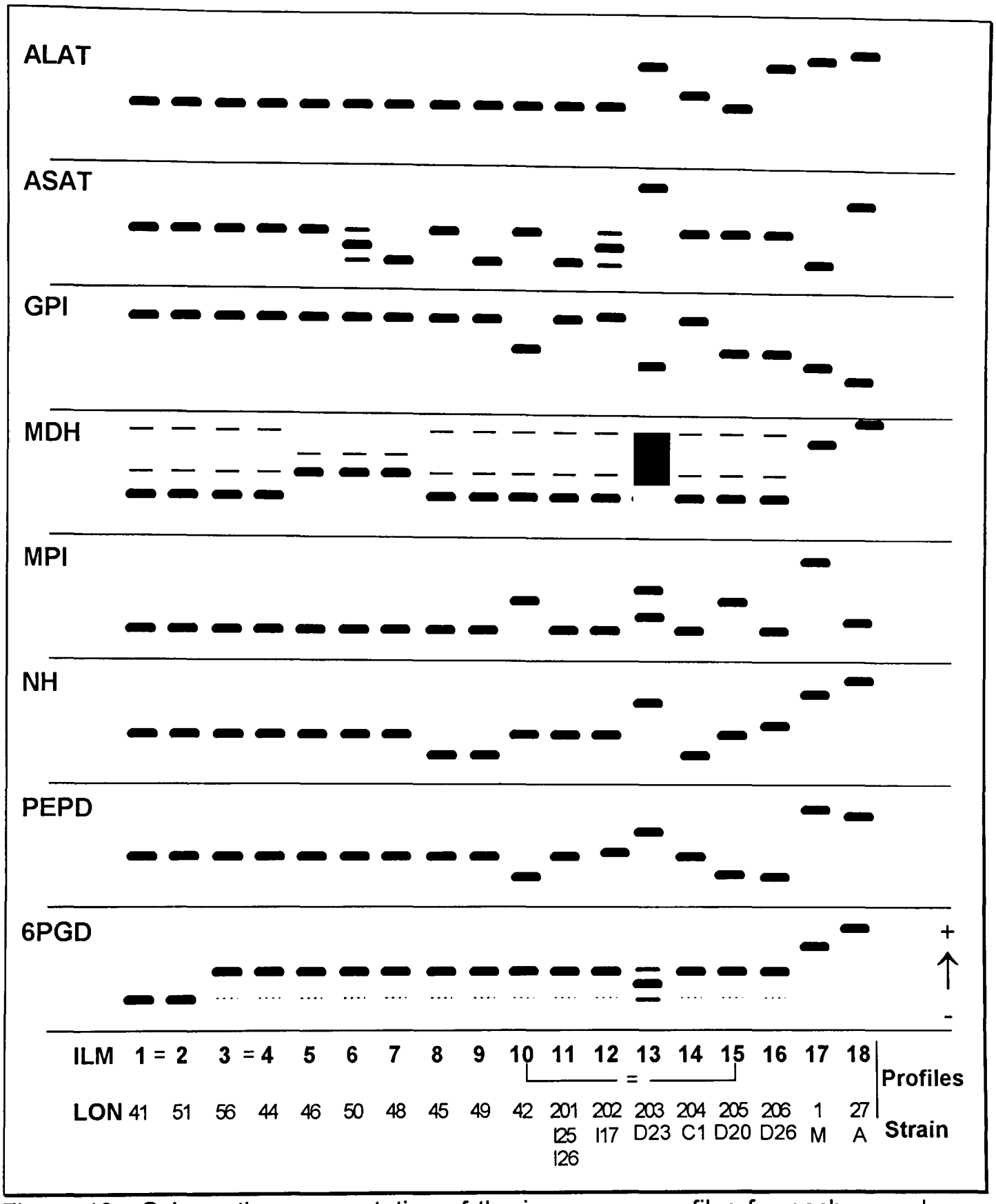


Figure 13 - Schematic representation of the isoenzyme profiles for each zymodeme type found. Not drawn to scale. Dashed lines indicate bands which presence was not reliable in different replicates but which had been used by Le Blancq (1986). Identical zymodeme profiles are indicated (=), which would be differentiated if MDH had been scored as by Le Blancq (1986). Bottom lines represent origin of migration (anode).

3.2.3. Phylogenetic analysis

The isoenzyme data obtained in this work was scored either as presence / absence characters or as qualitative multistate characters (each band coded as a base or an amino acid in, respectively, DNA or protein sequence files, and being analysed with the simplest algorithms by the appropriate programmes); 10.2. Annex 2 - Table XXXVI .

The shortest parsimony trees from binary coded data were produced by Dollo parsimony (Fig. 14). Rooting of those trees depended on the outgroup used: if ILM 13 was used as outgroup, then the root lay in the branch leading to ILM 10 and ILM 16, but if only ILM 17 and 18 were used as outgroups, then the root lay in the branch leading to ILM 8 and 9. Topology of the two trees was otherwise similar, except for position of ILM 11.

An UPGMA tree using Jaccard dissimilarity data (Table XVI) was rooted as the equivalent tree obtained by Le Blancq *et al.* (1986) using LON zymodemes (Fig. 15), but further topology within the *L. donovani* complex was different, except for clustering of ILM 8 with 9, and ILM 5 with 7. The UPGMA phenogram, however, was not affected by the outgroup (ILM 13, 17, 18) as the parsimony trees had been. In trees by neighbour-joining and Fitch-Margoliash (sum of squares (SQ) = 2.1; average percent standard deviation (ASD) = 14), ILM 3 clustered with ILM 1, and ILM 11 with ILM 5 and 7. However, ILM 11/5/7 were associated with ILM 8 and 9 or with ILM 1 and 3 by Fitch-Margoliash or by neighbour-joining clustering, respectively. Clustering methods using distances from data coded as amino acid residues in a protein sequence (Table XVI) produced further alternative trees.

Parsimony trees from data coded as bases in a DNA sequence varied enormously according to the rooting and in each analysis, therefore, a maximum likelihood tree is presented instead (Fig. 16), with a transition / transversion ratio set to 1. This tree is equivalent to that obtained by Rioux *et al.* (1990) except that ILM 5/6/7 were not in the root of the *L. donovani* complex, but branched from ILM 11, which is depicted here as part of a circle including ILM 3 and ILM 12 (a putative hybrid, heterozygous for ASAT). Another circle links ILM 5/6/7, in which ILM 6 is the putative hybrid, heterozygous for ASAT. Cladograms using MON data (**B** by Rioux *et al.* (1990) and **C**, coded as bases in a DNA sequence) were compared. Whilst general topologies are similar, the clear cut division between *L. donovani* and *L. infantum* seen in Rioux's cladogram is less obvious in the alternative parsimony tree and even less in the ILM data tree, where *L. infantum* is much less diverse than *L. donovani*.

Table XVI - Dissimilarity calculated between all zymodemes analysed: bottom half for binary coding using Jaccard index and upper half for coding as multistate qualitative characters (amino acids in a protein sequence like file) using Kimura parameter.

	ILM 1	ILM 3	ILM 5	ILM 7	ILM 8	ILM 9	ILM 10	ILM 11	ILM 13	ILM 16	ILM 17
ILM 1	-	0.000	0.160	0.360	0.159	0.356	0.626	0.159	-1.000	1.013	-1.000
ILM 3	0.222	-	0.137	0.304	0.137	0.304	0.516	0.137	-1.000	0.799	-1.000
ILM 5	0.400	0.222	-	0.137	0.304	0.516	0.799	0.304	-1.000	1.214	-1.000
ILM 7	0.545	0.400	0.222	-	0.516	0.304	1.214	0.137	1.984	1.984	-1.000
ILM 8	0.400	0.222	0.400	0.545	-	0.137	0.799	0.304	-1.000	1.214	-1.000
ILM 9	0.545	0.400	0.545	0.400	0.222	-	1.214	0.137	1.984	1.984	-1.000
ILM 10	0.667	0.545	0.667	0.769	0.667	0.769	-	0.799	-1.000	0.137	-1.000
ILM 11	0.400	0.222	0.400	0.222	0.400	0.222	0.667	-	1.984	1.214	-1.000
ILM 13	0.941	0.941	0.941	0.941	0.941	0.941	0.941	0.941	-	1.984	-1.000
ILM 16	0.769	0.667	0.769	0.857	0.667	0.769	0.545	0.769	0.875	-	-1.000
ILM 17	1.000	1.000	1.000	0.933	1.000	0.933	1.000	0.933	0.941	1.000	-
ILM 18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

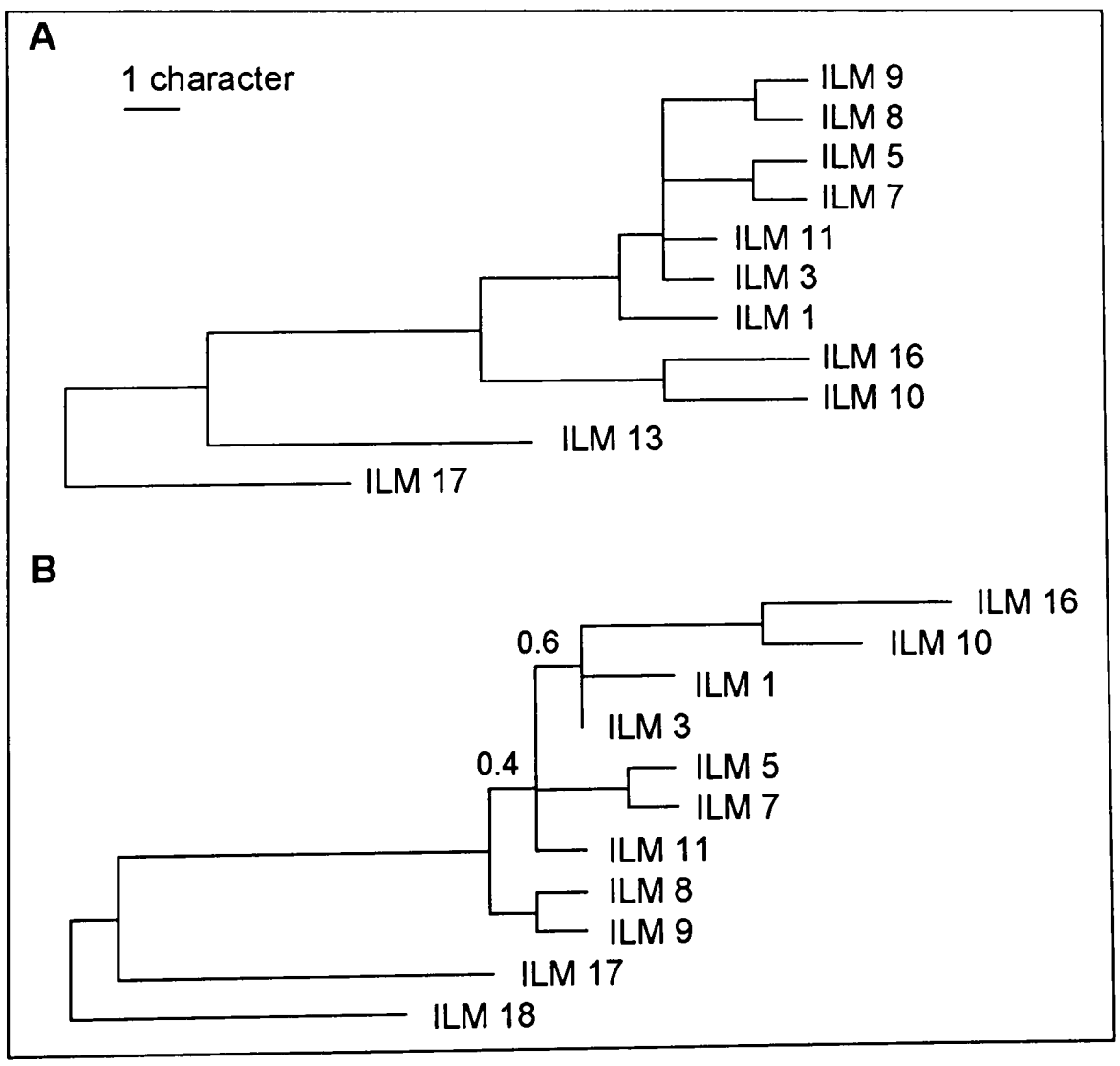


Figure 14 - Parsimony analysis of the isoenzyme profiles determined with eight enzymes. Consensus trees from Dollo parsimony, using binary coded data: A) with ILM 17 and 13 as outgroups and from 18 best trees with 20 reversions in each character; B) with ILM 17 and 18 as outgroups and from 42 trees with 13 reversions in each character. All branches were present in more than eight out of 10 trees, except for those which have relative frequencies indicated above the branch.

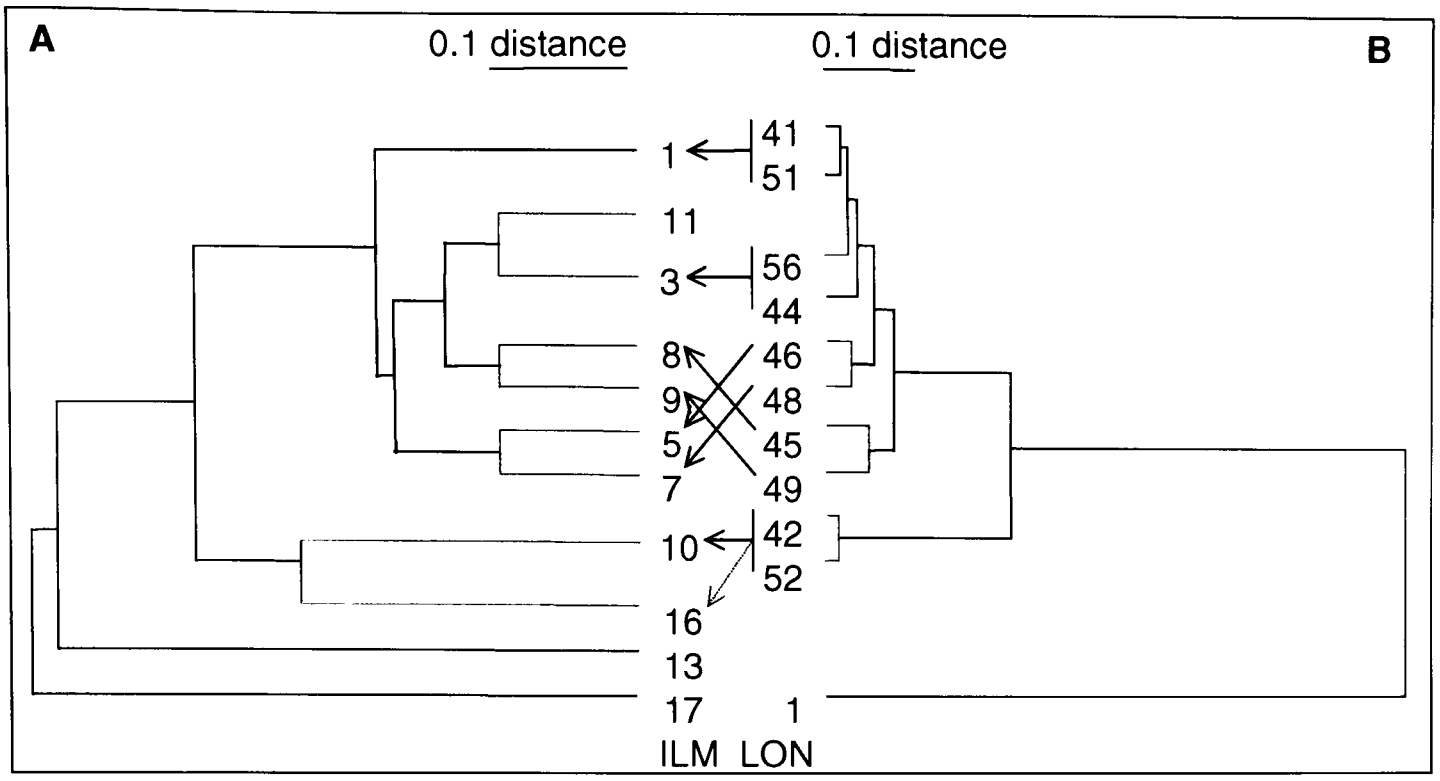


Figure 15 - Comparative UPGMA phylogenetic analysis of the isoenzyme profiles determined in this work (eight enzymes) with that of previous workers (12 enzymes). A) UPGMA tree from Jaccard dissimilarity compared with B) UPGMA tree obtained by Le Blancq *et al.* (1986). Correspondence between zymodemes is shown by arrows. A broken line indicates correspondence of strains, but not of zymodemes (see text).

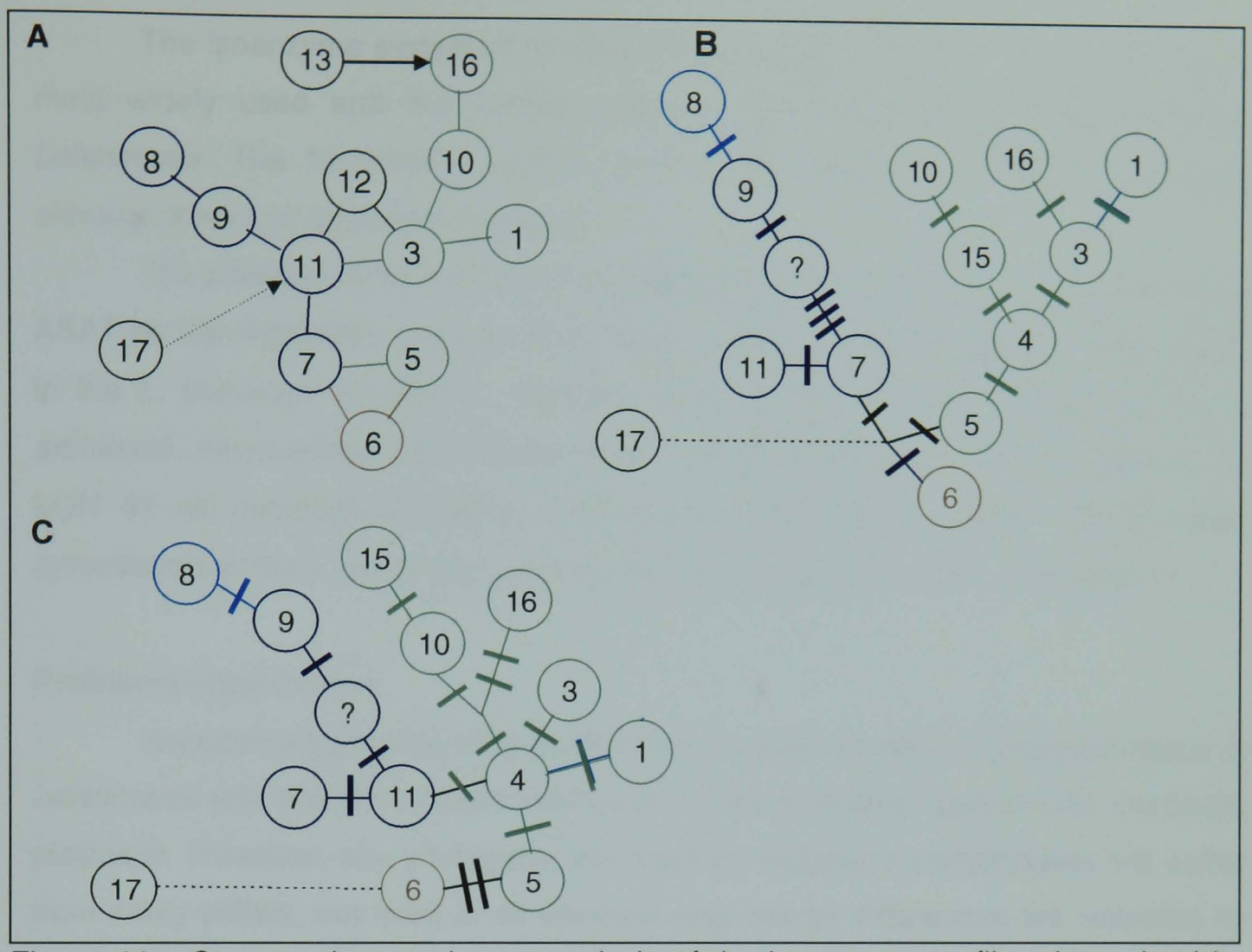


Figure 16 - Comparative parsimony analysis of the isoenzyme profiles determined in this work with that of previous workers. Comparison of A) maximum likelihood cladogram (without branch lengths) obtained with data from the present work and scored as multistate qualitative data in the form of bases in a DNA sequence (using transition / transversion ratio = 1) with B) a cladogram obtained by Rioux *et al.* (1990) and C) a reconstruction of a MON IEA parsimony tree with data (same as in B) scored as DNA sequence data (as in A) (consensus from 100 most parsimonious trees). ILM zymodemes are represented by numbers inside circles and MON codes are represented in trees B and C by ILM zymodeme numbers: MON 1 is 9, MON 24 is 11, MON 18 is 5, MON 30 is 7, MON 31 is 10, MON 32 is 3, MON 35 is 16, MON 37 is 4, MON 82 is 6, MON 83 is 15, MON 34 is ?. In blue are *L. infantum* strains, as defined by Rioux *et al.* (1990), in green are *L. donovani* strains, in red is *L. archibaldi* and in CYAN is zymodeme ILM 8, which was used in this work but has no known correspondence in MON. ILM 6 and ILM 12, heterozygous for ASAT, were not included in the ILM analysis but were placed here as putative ILM 5 and ILM 7 or ILM 11 and ILM 3 hybrid, respectively, hence the triangular branching patterns on A). Alternative root position using ILM 13 as outgroup is shown in A. The number of steps for each branch is represented by dashes in B and C.

3.3. Discussion

The isoenzyme system of 15 enzymes described by Rioux *et al.* (1990) is the most widely used and the current reference for typing and taxonomy of OW *Leishmania*. The 12 enzyme system used in the LSHTM has fallen into disuse, although it was an important reference.

The cladistic analysis of MON zymodeme profiles by Rioux *et al.* (1990) used ASAT as the diagnostic enzyme (Rioux *et al.*, 1990) to differentiate two main clades in the *L. donovani* complex *L. infantum* and *L. donovani*, and a third species *L. archibaldi*. Alternatively, LON 49 has been used as synonymous with *L. infantum ss*, LON 41 as synonymous with *L. donovani ss* (Le Blancq, 1986), and all other zymodemes in the *L. donovani* complex have been designated as *L. donovani sl*.

Problems encountered

Isoenzyme typing has historically been the gold standard in characterization of *Leishmania* and any DNA based method has to be evaluated against it for continuity purposes. However, any phylogeny produced by enzyme electrophoresis will suffer from many pitfalls, not least of all because only charge differences are detected by this method and many changes in amino acid composition are missed. Other problems are the low number of characters produced, the impossibility of correlating band size with similarity and the dependence on growing conditions for expression of determined enzymes or alleles.

The profiles observed for some strains were not those expected from previous characterizations by Le Blancq or Evans. Some differences were strain restricted. In the case of strain Sukkar 2 (D23), the profiles were so different from all other strains used here, that the possibility of an error retrieving the right strain is very strong. In the case of strain Wangjie 1, it had been typed as LON 42, but in here a small difference was detected in size of the NH band, the size of which was equivalent to that described by Rioux *et al.* (1990). This strain was also found here to have different profiles for enzymes ALAT, MPI and PEPD than those described by Le Blancq (1986). Some *L. infantum* strains previously typed as LON 49 (Le Blancq, 1986) were found here to have different profiles - L82 (I1) with a *L. donovani* type ASAT, and Buck (I31) with the alternative NH pattern. Other differences were generic for one enzyme (MDH, ALAT and 6PGD), for which some bands were difficult to discriminate by size (MDH) and others were faint and not reproducible. These secondary bands are probably artifacts, most likely dependent on particular running conditions or enzyme extractions.

The reasons for differences in particular strain profiles are not clear, although cross contamination is a possibility in the original typing. Great care was taken thoughout the present work and particular DNA based profiles, most notably that of Wangjie 1, were consistent with their respective MON zymodemes.

There are several possible explanations for the difficulties encountered here in replicating certain profiles, especially regarding faint bands. Enzyme extractions could have been done from parasites in different stages of growth, or the culture medium could have been different from the one previously used, thus slightly altering enzyme expression. Running or staining conditions may not have been totally reproduced. However, only minimal changes would be expected to result from these latter variables, such as satellite bands and increasing or decreasing resolution.

The difficulties faced in this work to fully reproduce the profiles found by previous workers using the same system are a small example of the problems faced when trying to standardize the method for use in different laboratories. Not only the same reference strains are required, but diverse laboratories use different methodologies for culture, electrophoresis and development.

Zymodemes

An alternative classification to the LON was used in this work because of the differences between the profiles found now with those previously by Le Blancq (1986), the scoring method used and because enzymes that had not been found to be polymorphic in the *L. donovani* complex were not tested here. The new series uses the prefix ILM and describes 15 distinct zymodemes. Some strains, which had not been characterized before by the LON system, were found to have novel profiles. These strains were tentatively ascribed a LON number, starting from 200, exclusively and were described here with the respective profiles. Because further characterization is required stabilates were stored in the LSHTM.

Two zymodemes described here were found to be heterozygous for ASAT, thus suggesting that recombination may occur within the *L. donovani* complex. Heterozygosity may have arisen independently with descendants becoming homozygous, but such a scenario is not likely because heterozygosity would have to have arisen independently twice producing the same alleles.

The enzyme typing of the strains used in this project provided a basis for comparison of traditional typing methods with DNA based methods. Although a comprehensive allozyme study of the *L. donovani* complex was not achieved in this work the characterization of strains was sufficient to provide a correlation with established isoenzyme typing.

Phylogenies

The isoenzyme data collected here was used to generate phylogenies of the *L. donovani* complex in order to compare with phylogenies generated by DNA based methods. In all obtained phylogenies, there were apparent associations of zymodemes (ILM 5, 7 and 6, ILM 8 and 9, ILM 10 and 16) but topologies were variable, according to the coding method, outgroup and tree building method used. The simplification of the MDH data had as consequence a reduction in tree resolution and higher topological uncertainty. However, it was evident that the clear division of the *L. donovani* complex into *L. donovani*, *L. infantum* and *L. archibaldi,* as suggested by Rioux *et al.* (1990), was not well supported, inclusively by re-analysis of their data. Phylogenetic reconstructions based on isoenzyme characters, especially using a small number of characters are not robust and should be interpreted with care when addressing *Leishmania* taxonomy.

4. Comparison of *Leishmania* genomes: RAPD analysis

4.1. Introduction

A simple method to generate random anonymous genetic markers by PCR amplification was developed simultaneously by Williams *et al.* (1990) and Welsh and McClelland (1990) and was called random amplified polymorphic DNA (RAPD) or arbitrarily primed PCR (AP-PCR), respectively. A single primer of, usually, short length (8-12 bases) is used in a PCR reaction, frequently with low annealing temperatures during thermal cycling, and thus low specificity. Potentially, a number of fragments are amplified from undetermined regions in the genome. Whole genomes can be compared without the need to have previous information on DNA sequence, through this fingerprinting method. Lack of reproducibility due to sensitivity to experimental conditions is a common problem with RAPDs (Lamboy, 1994; Noyes *et al.*, 1996) but it has been a useful technique in producing taxonomic data for *Leishmania* (Andresen *et al.*, 1996; Motazedian *et al.*, 1996; Noyes *et al.*, 1996). By careful choice of primers the taxonomic discrimination level can be manipulated, although RAPD is usually applicable to strain and species levels.

A fragment is amplified if the primer sequence is present in the 5' terminus of both ends of the target sequence, so that a single point mutation can produce or eliminate a band if located in either priming site. Deletions and insertions can also modify RAPD patterns and it is not possible to identify homologous bands except by probing blots with labelled isolated fragments, or by comparative DNA sequencing. An amplified character is dominant over a non-amplification, therefore it is not possible to determine homozygosity or heterozygosity for each amplified band.

In this work, RAPDs were used to compare: *L. infantum* and *L. chagasi*; Portuguese strains of *L. infantum,* and *L. donovani* strains.

4.2. Quality control of RAPD amplification

The effect of DNA concentration on RAPD profiles was assessed for all primers but A5, that is, a total of nine primers in all, six of which are shown in Figure 17. The three *L. donovani* complex reference strains - DD8, PP75 and IPT-1 - were used with four different DNA concentrations (0.05, 0.5, 1.25, 5 ng/μl).

Profiles were generally reproducible across the range of DNA concentrations tested. In some cases, the most extreme concentrations - 0.05 and mostly 5ng/μl - did not produce reliable profiles. The DNA concentration selected was 1.25ng/μl.

Lower DNA concentrations worked equaly well and small variations in DNA concentration did not greatly affect results. Relative intensity of bands, especially shorter fragments in relation to longer fragments, could vary somewhat with DNA concentration, although not in the middle concentration values tested.

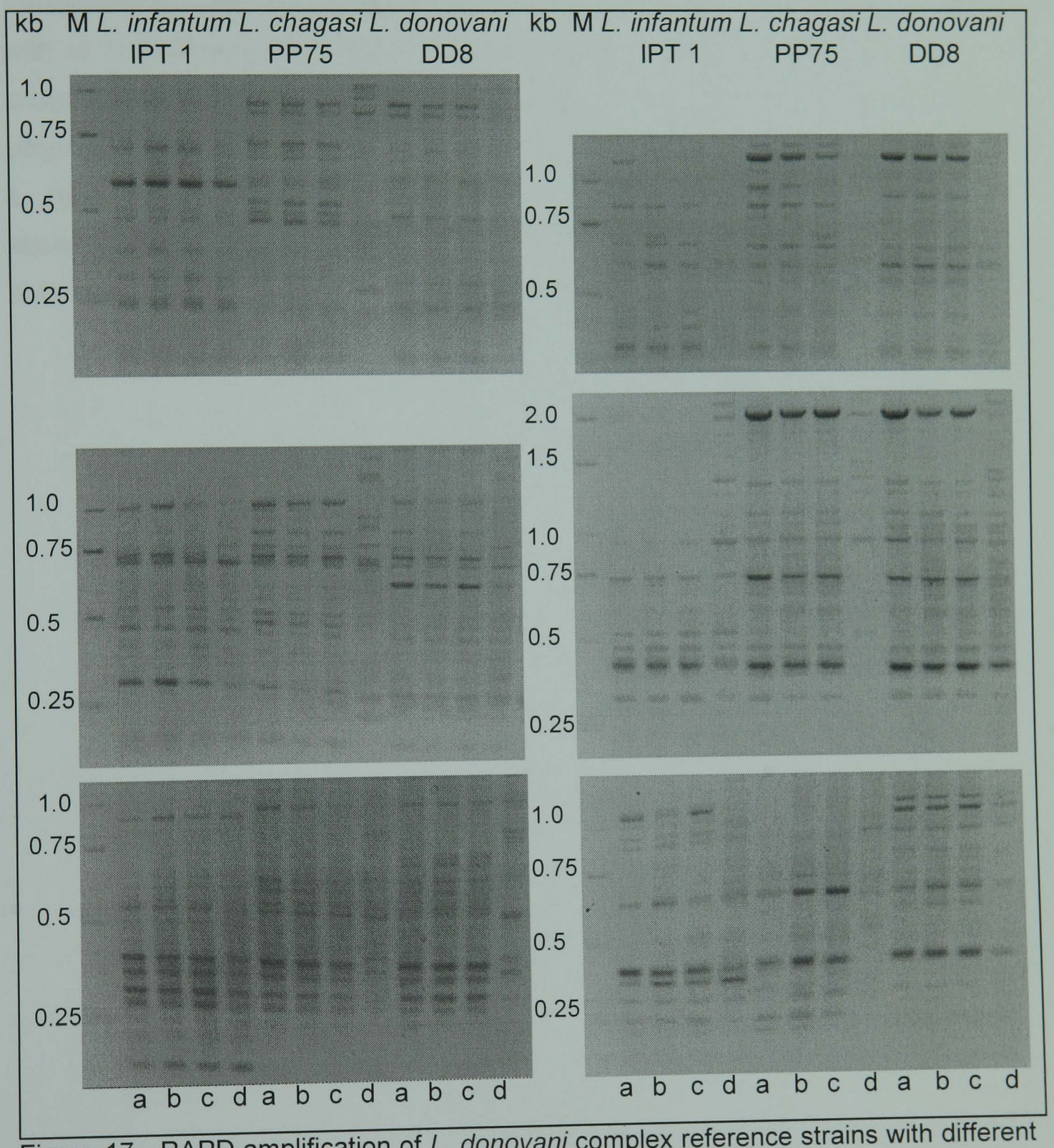


Figure 17 - RAPD amplification of *L. donovani* complex reference strains with different DNA concentrations (a, b, c, d; respectively 0.05, 0.5, 1.25 and 5ng/μl), for primers (left to right, top to bottom) D3, D8, D10, A6, A4, A5. 'M' are molecular weight markers.

4.3. The relationship between *Leishmania infantum* and *Leishmania chagasi* and their relation to the *Leishmania donovani* complex.

RAPD amplification, using primers A2, A4, A5, A6, D3, D8, D10, H1, H4 and L2, produced very distinct profiles between the selected strains of Old World *Leishmania* species complexes: *L. donovani*, *L. aethiopica*, *L. major* and *L. tropica*, with all the primers tested. Furthermore, primer H1 separated the *L. infantum* / *L. chagasi* strains from the *L. donovani* strains (Fig. 18A). Genetic diversity detected as polymorphic RAPD profiles, was identified within the *L. donovani* complex and within *L. infantum* / *L. chagasi* (Fig. 18B). RAPD bands were scored for presence or absence (Annex 2) and genetic distances (Table XVII) were calculated to produce phenograms.

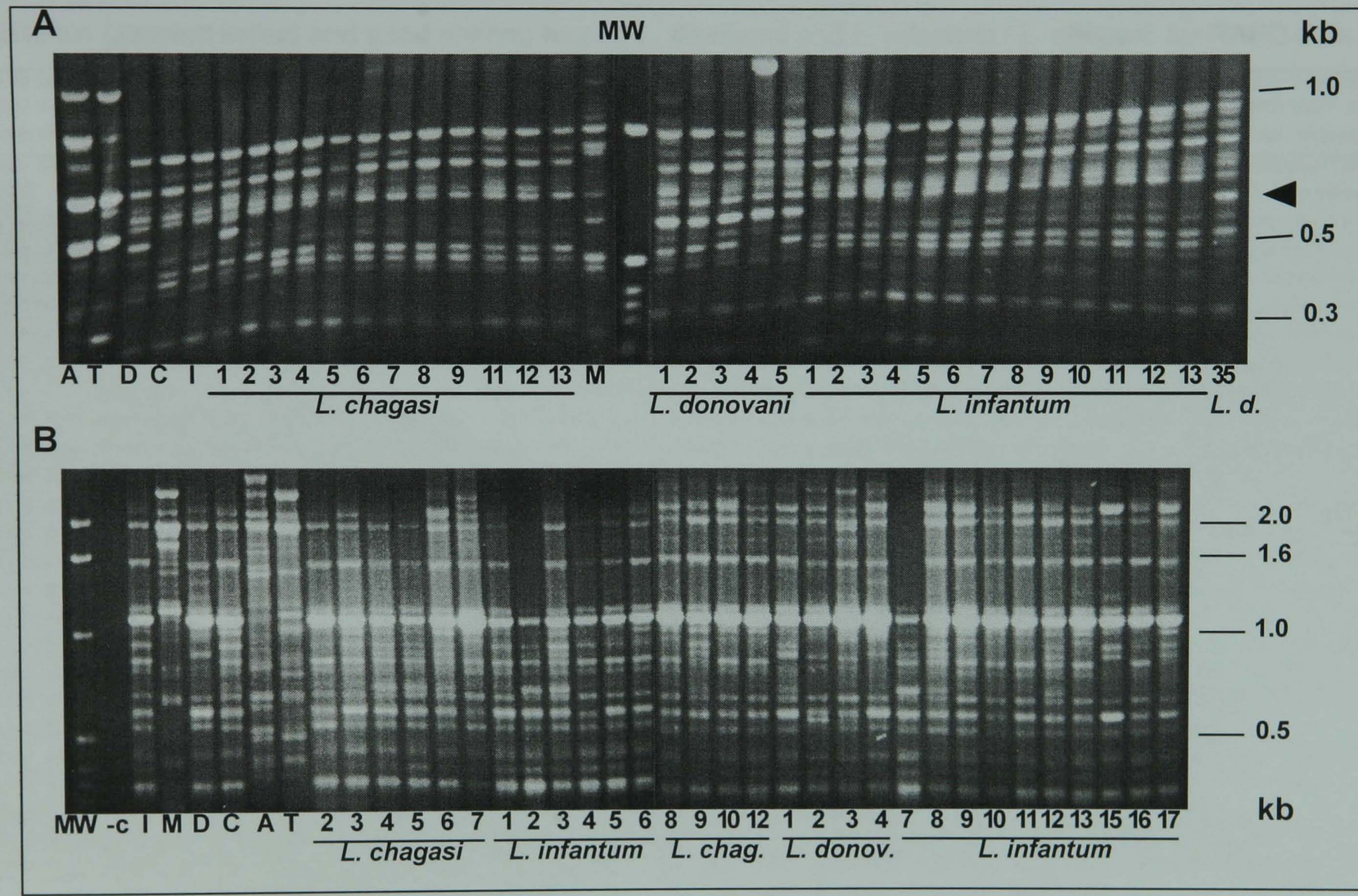


Figure 18 - RAPD profiles for strains of the *L. donovani* complex obtained with primers H1 (A) and D8 (B). The band shown by the arrow discriminated between *L. infantum* / *L. chagasi* and *L. donovani* strains and D8 generated polymorphic profiles within *L. infantum* and *L. chagasi*. Reference strains: A - *L. aethiopica*; C - *L. chagasi*; I - *L. infantum*; D - *L. donovani*; T - *L. tropica*. For strain codes, see Tables II to V. -c is negative control.

Table XVII. Pairwise distance (Jaccard index) and band sharing among *L. donovani* and *L. infantum / L. chagasi* for RAPD data.

	5-ASKH	L96	K27	IPT-1	L82	Pharoah	Lombardi	LEM75	IMPT104	IMT150	IMT152	IMT89	IMT124	IMT108	IMT193	Strain A	L53	PP75*	WR285*	CO910*	M9702*	M8270*	M12727*	M12734*	M12337*	M7633*	M12085*	M12084*	DD8	HU3	MRC(L)3	MRC74	Patna1
5-ASKH	82	20/156	24/181	24/185	23/178	25/188	22/169	24/180	23/184	24/189	23/185	24/188	23/182	24/189	22/185	24/181	25/183	23/181	23/183	23/184	22/180	22/171	24/188	24/184	23/182	25/189	22/181	24/185	22/184	24/180	26/187	22/183	26/188
L96	0.853	74	44/173	21/177	20/170	22/180	19/161	21/172	21/176	21/181	21/177	21/180	21/174	21/181	20/177	21/173	22/175	21/173	20/175	20/176	20/172	18/163	21/180	22/176	19/174	23/181	19/173	21/177	20/176	22/172	21/179	20/175	21/180
K27	0.847	0.659	99	25/202	25/195	24/205	19/186	25/197	24/201	24/206	24/202	24/205	24/199	25/206	24/202	24/198	25/200	25/198	24/200	25/201	25/197	20/188	24/205	24/201	24/199	26/206	24/198	24/202	25/201	25/197	24/204	23/200	25/205
IPT-1	0.851	0.865	0.859	103	94/199	97/209	84/190	94/201	96/205	99/210	96/206	97/209	93/203	99/210	95/206	92/202	92/204	95/202	94/204	96/205	94/201	83/192	97/209	93/205	94/203	96/210	93/202	96/206	86/205	83/201	83/208	78/204	83/209
L82	0.852	0.867	0.853	0.105	96	93/202	84/183	93/194	93/198	93/203	92/199	93/202	90/196	94/203	91/199	88/195	89/197	90/195	92/197	94/198	92/194	84/185	93/202	90/198	91/196	92/203	88/195	92/199	81/198	82/194	78/201	74/197	77/202
Pharoah	0.847	0.861	0.867	0.134	0.147	106	85/193	96/204	101/208	104/213	100/209	102/212	98/206	104/213	98/209	97/205	98/207	97/205	95/207	96/208	95/204	87/195	102/212	97/208	96/206	101/213	94/205	99/209	87/208	87/204	84/211	80/207	86/212
Lombardi	0.850	0.866	0.886	0.208	0.152	0.213	87	84/185	83/189	84/194	82/190	83/193	81/187	84/194	81/190	82/186	80/188	81/186	86/188	83/189	81/185	79/176	83/193	81/189	81/187	82/194	78/186	82/190	72/189	75/185	71/192	69/188	70/193
LEM75	0.846	0.861	0.855	0.121	0.079	0.111	0.168	98	96/200	96/205	95/201	96/204	93/198	96/205	93/201	91/197	92/199	91/197	93/199	94/200	91/196	83/187	96/204	93/200	93/198	95/205	89/197	94/201	82/200	86/196	80/203	76/199	81/204
IMPT104	0.857	0.865	0.864	0.119	0.114	0.056	0.217	0.077	102	102/209	101/205	101/208	99/202	100/209	98/205	95/201	96/203	95/201	93/203	96/204	93/200	83/191	100/208	97/204	96/202	99/209	94/201	98/205	86/204	85/200	83/207	79/203	84/208
IMT150	0.855	0.869	0.868	0.108	0.155	0.046	0.236	0.119	0.047	107	102/210	102/213	99/207	103/214	99/210	96/206	98/208	97/206	94/208	96/209	95/205	85/196	103/213	97/209	96/207	102/214	95/206	99/210	88/209	86/205	84/212	80/208	87/213
IMT152	0.858	0.865	0.865	0.127	0.140	0.083	0.241	0.104	0.029	0.056	103	102/209	100/203	99/210	99/206	96/202	96/204	95/202	92/204	95/205	92/201	82/192	100/209	98/205	97/203	99/210	96/202	99/206	87/205	84/201	84/208	80/204	86/209
IMT89	0.854	0.868	0.867	0.134	0.147	0.073	0.245	0.111	0.056	0.081	0.047	106	100/206	103/213	102/209	96/205	97/207	95/205	94/207	97/208	92/204	83/195	101/212	100/208	99/206	100/213	98/205	103/209	87/208	86/204	86/211	81/207	88/212
IMT124	0.855	0.863	0.863	0.155	0.151	0.093	0.236	0.114	0.039	0.083	0.029	0.057	100	97/207	97/203	95/199	95/201	94/199	90/201	93/202	90/198	82/189	97/206	98/202	95/200	96/207	94/199	97/203	85/202	84/198	83/205	80/201	84/206
IMT108	0.855	0.869	0.862	0.108	0.138	0.046	0.236	0.119	0.083	0.072	0.108	0.064	0.118	107	100/210	96/206	98/208	98/206	96/208	98/209	96/205	86/196	102/213	97/209	97/207	101/214	97/206	100/210	89/209	87/205	83/212	78/208	88/213
IMT193	0.865	0.873	0.865	0.144	0.157	0.117	0.257	0.139	0.084	0.108	0.075	0.047	0.085	0.091	103	93/202	93/204	92/202	92/204	95/205	90/201	80/192	97/209	96/205	96/203	96/210	95/202	99/206	86/205	82/201	83/208	80/204	84/209
Strain A	0.847	0.862	0.862	0.164	0.178	0.102	0.212	0.142	0.104	0.127	0.094	0.119	0.087	0.127	0.147	99	94/200	94/198	90/200	91/201	88/197	82/188	95/205	94/201	92/199	94/206	92/198	93/202	84/201	84/197	82/204	80/200	83/205
L53	0.842	0.856	0.857	0.179	0.176	0.101	0.259	0.140	0.103	0.109	0.111	0.118	0.104	0.109	0.162	0.113	101	94/200	90/202	93/203	90/199	83/190	98/207	95/203	92/201	97/208	92/200	94/204	85/203	86/199	80/206	77/202	86/207
PP75*	0.854	0.862	0.855	0.112	0.143	0.102	0.229	0.142	0.104	0.110	0.112	0.136	0.105	0.093	0.164	0.096	0.113	99	91/200	94/201	92/197	84/188	98/205	95/201	93/199	97/206	93/198	94/202	84/201	83/197	80/204	77/200	82/205
WR285*	0.856	0.871	0.864	0.145	0.124	0.152	0.157	0.123	0.155	0.175	0.179	0.168	0.189	0.143	0.179	0.182	0.196	0.165	101	96/203	93/199	83/190	95/207	91/203	92/201	94/208	89/200	93/204	82/203	83/199	83/206	77/202	82/207
CO910*	0.857	0.872	0.858	0.119	0.096	0.143	0.217	0.113	0.111	0.150	0.136	0.126	0.147	0.117	0.136	0.173	0.155	0.121	0.103	102	94/200	83/191	97/208	94/204	95/202	96/209	92/201	96/205	84/204	83/200	82/207	76/203	82/208
M9702*	0.861	0.868	0.855	0.121	0.098	0.128	0.221	0.133	0.131	0.136	0.156	0.179	0.167	0.119	0.189	0.193	0.174	0.124	0.123	0.113	98	85/187	95/204	90/200	91/198	94/205	89/197	91/201	82/200	81/196	78/203	74/199	80/204
M8270*	0.852	0.876	0.881	0.239	0.168	0.194	0.186	0.202	0.231	0.234	0.255	0.259	0.234	0.218	0.286	0.226	0.224	0.192	0.224	0.231	0.167	89	86/195	83/191	82/189	85/196	80/188	82/192	73/191	75/187	70/194	67/190	72/195
M12727*	0.854	0.868	0.867	0.134	0.147	0.073	0.245	0.111	0.074	0.064	0.083	0.090	0.110	0.081	0.134	0.136	0.101	0.084	0.152	0.126	0.128	0.211	106	100/208	99/206	104/213	96/205	100/209	86/208	87/204	84/211	80/207	86/212
M12734*	0.850	0.857	0.864	0.170	0.167	0.126	0.250	0.131	0.093	0.134	0.084	0.074	0.058	0.134	0.119	0.121	0.120	0.104	0.188	0.145	0.182	0.231	0.074	102	98/202	98/209	95/201	99/205	83/204	85/200	84/207	81/203	84/208
M12337*	0.855	0.877	0.863	0.138	0.135	0.127	0.236	0.114	0.094	0.135	0.085	0.075	0.095	0.118	0.103	0.140	0.156	0.123	0.156	0.112	0.150	0.234	0.075	0.058	100	98/207	95/199	99/203	83/202	83/198	83/205	79/201	82/206
M7633*	0.848	0.854	0.856	0.158	0.171	0.098	0.268	0.136	0.100	0.089	0.108	0.115	0.135	0.106	0.158	0.161	0.126	0.110	0.175	0.150	0.153	0.234	0.046	0.117	0.101	107	96/206	100/210	85/209	86/205	84/212	81/208	85/213
M12085*	0.862	0.877	0.862	0.147	0.178	0.153	0.278	0.176	0.121	0.144	0.094	0.084	0.105	0.110	0.112	0.132	0.148	0.114	0.198	0.156	0.176	0.259	0.119	0.104	0.087	0.127	99	97/202	84/201	80/197	81/204	76/200	86/205
M12084*	0.851	0.865	0.865	0.127	0.140	0.100	0.241	0.121	0.084	0.108	0.075	**0.028**	0.085	0.091	0.075	0.147	0.145	0.130	0.162	0.119	0.173	0.255	0.083	0.086	0.048	0.091	0.076	103	84/205	84/201	85/208	80/204	85/209
DD8	0.864	0.872	0.858	0.277	0.308	0.281	0.385	0.305	0.271	0.273	0.263	0.281	0.274	0.258	0.277	0.282	0.280	0.282	0.322	0.300	0.305	0.381	0.295	0.314	0.303	0.315	0.282	0.306	102	83/200	83/207	78/203	91/208
HU3	0.846	0.853	0.855	0.297	0.268	0.256	0.318	0.218	0.261	0.277	0.282	0.271	0.263	0.263	0.311	0.257	0.239	0.272	0.284	0.291	0.296	0.330	0.256	0.261	0.278	0.277	0.316	0.282	0.291	98	83/203	80/199	84/204
MRC(L)3	0.839	0.867	0.867	0.336	0.366	0.339	0.413	0.350	0.331	0.344	0.323	0.312	0.320	0.357	0.336	0.328	0.365	0.355	0.325	0.344	0.376	0.435	0.339	0.317	0.320	0.344	0.341	0.309	0.331	0.308	105	93/206	87/211
MRC74	0.863	0.871	0.870	0.381	0.398	0.370	0.420	0.382	0.363	0.375	0.355	0.357	0.339	0.400	0.355	0.333	0.384	0.374	0.384	0.402	0.408	**0.455**	0.370	0.336	0.352	0.362	0.387	0.355	0.376	0.328	0.177	101	78/207
Patna1	0.840	0.868	0.861	0.341	0.384	0.317	0.431	0.341	0.323	0.310	0.301	0.290	0.311	0.296	0.328	0.320	0.289	0.333	0.344	0.349	0.355	0.415	0.317	0.323	0.339	0.336	0.277	0.315	0.222	0.300	0.298	0.395	106

The three groups of strains are, from top: outgroup , *L. infantum / L. chagasi** and *L. donovani*. In the upper triangular matrix are shared over total number of bands of each strain pair. In the lower matrix are Jaccard distances and in the diagonal the number of bands for each strain. (Continues)

(Table XVII) Bold underlined are maximum and minimum distances within the *L. donovani* complex. Distances smaller than 0.273 (60% of the maximum distance within the *L. donovani* complex) are shaded.

In both single linkage and UPGMA dendrograms, built from a Jaccard distance matrix (Table XVII), *L. infantum* and *L. chagasi* were resolved from the *L. donovani* strains but not from each other (Fig. 19), forming a single branch within *L. donovani. Leishmania donovani* strains of the same geographical origin (Indian or Kenyan) were grouped. The dendrograms were rooted to the outgroups (*L. aethiopica*, *L. major*, *L. tropica*) between the *L. donovani* Kenyan strains and the remainder. Other distance coefficients, like simple matching [1-(a+d)/n], Yule [1-(ad-bc)/(ad+bc)] and Euclidean distance [$\sqrt{(b+c)}$] produced topologies similar to single linkage and UPGMA. The cophenetic correlation coefficient (CC) was high using either method (0.994 for single linkage and 0.996 for UPGMA), suggesting that the dendrograms were good representations of the patristic distances, however, it was not possible to measure robustness of the nodes by bootstrap analysis. A three dimensional principal coordinates analysis, plotted with a superimposed minimum spanning tree (Fig. 20), illustrated again that *L. infantum* / *L. chagasi* were indistinguishable, and that *L. donovani* strains grouped according to geographical origin.

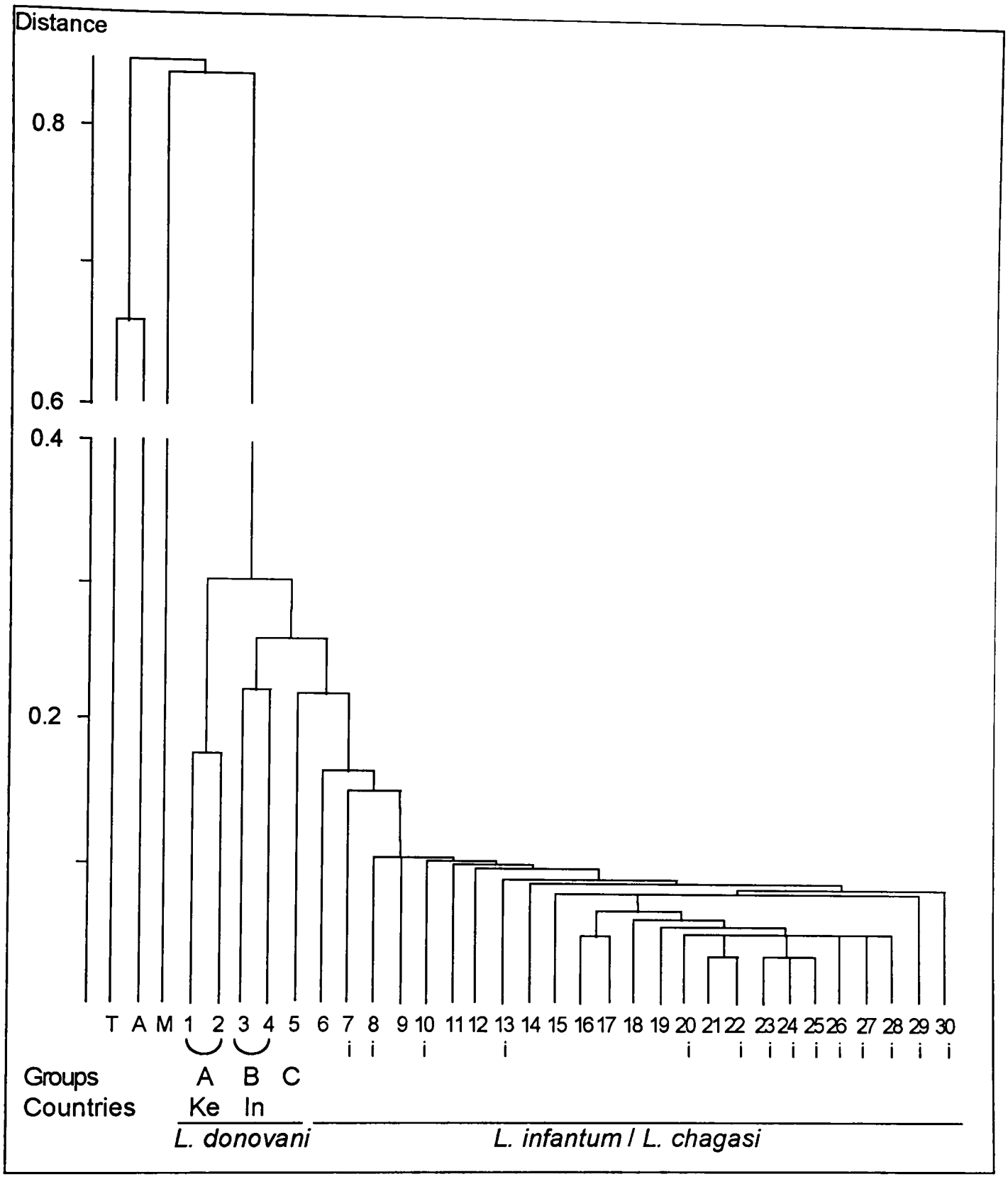


Figure 19. Single linkage dendrogram built from Jaccard distances recorded for RAPD data in strains of the *L. donovani* complex. Outgroups are *L. tropica* (**T**, L96), *L. aethiopica* (**A**, K27) and *L. major* (**M**, 5ASKH) strains. CC = 0.994. *Leishmania donovani* (1-5): 1- MRC(L)3; 2- MRC74; 3- DD8; 4- Patna 1; 5 - HU3, **Ke** - Kenyan; **In** - Indian. *Leishmania infantum* ('i', dark green) strains are mixed with *L. chagasi* ('c', light green) (6-30): 6- M8270; 7- Lombardi; 8- IPT-1; 9- WR285; 10- L53; 11- M9702; 12- CO910; 13- Strain A; 14- PP75; 15- M12085; 16- DOG 124; 17- M7633; 18- DOG 136; 19- DOG 118; 20- IMT193; 21- M12084; 22- IMT89; 23- IMPT104; 24- IMT152; 25- IMT124; 26- IMT108; 27- IMT150; 28- Pharoah; 29- LEM75; 30- L82. *Leishmania donovani* groups are as in Table XXXII (insert).

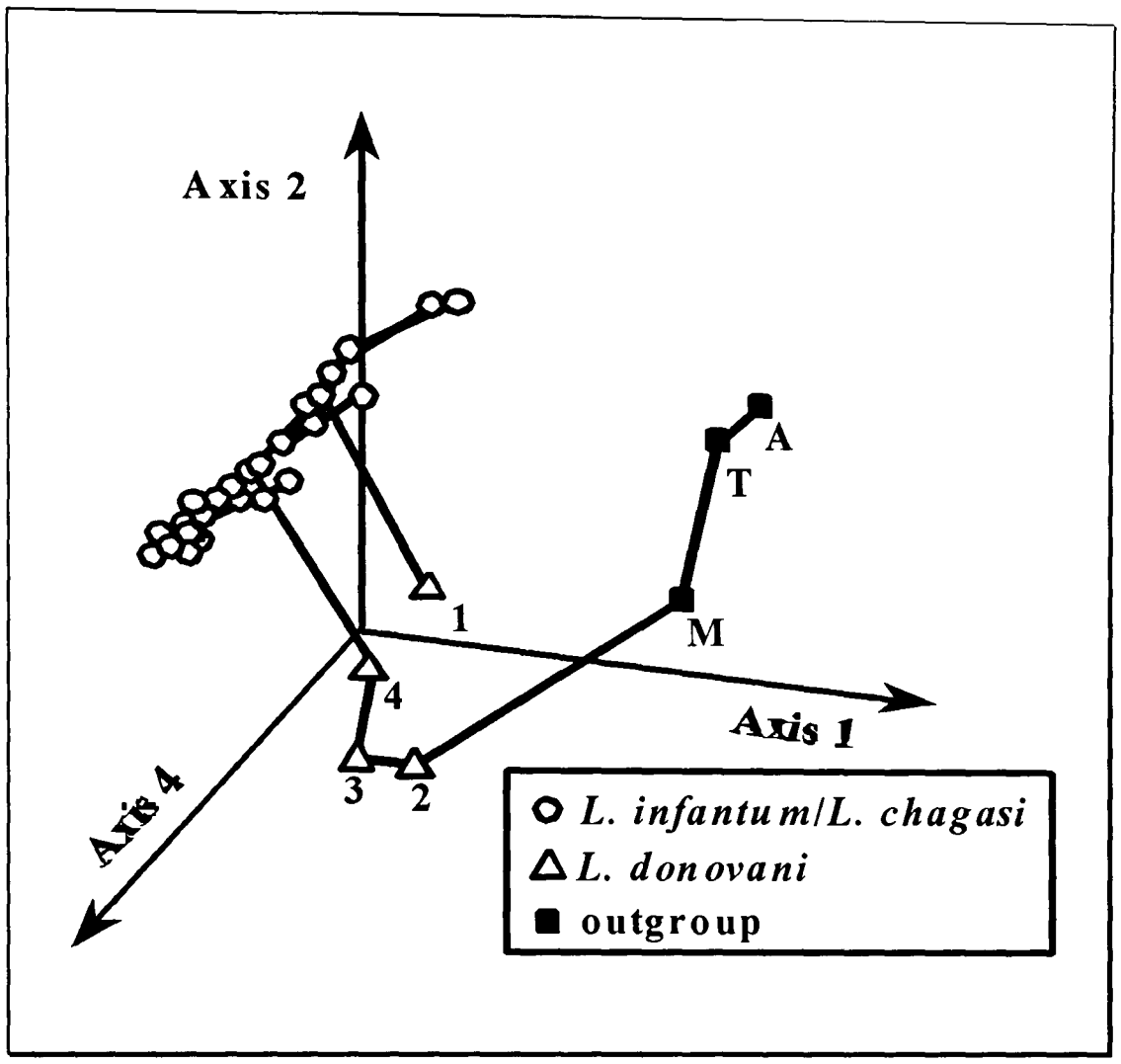


Figure 20 - A principal coordinates analysis with a superimposed minimum spanning tree depicting the first (29.48%), second (10.69%) and fourth axes as coordinates (cumulative percentage: 47.25%), from RAPD data. Outgrouped by *L. aethiopica* (A), *L. major* (M) and *L. tropica* (T) strains. Note that strains of *L. infantum / L. chagasi* are separated from *L. donovani* and that the latter show geographical patterns. *Leishmania donovani* strains are: 1 (group C) - HU3; Kenyan, 2 (group A)- MRC(L)3 and MRC74; Indian (group B), 3- Patna 1, 4- DD8. *Leishmania donovani* groups are as in Table XXXII (insert).

4.4. Genetic characterization of Portuguese *Leishmania*

All available Portuguese strains, from the four main foci were studied by RAPD analysis using the same ten primers as previously: A2, A4, A5, A6, D3, D8, D10, H1, H4 and L2. Other *L. infantum* and *L. chagasi* strains were used for comparisons, and *L. donovani*, *L. major* and *L. aethiopica* reference strains were used as outgroups. The analysis included all 22 Portuguese *L. infantum* strains, although some strains did not produce good quality profiles, and thus, only 18 Portuguese *L. infantum* strains were chosen for subsequent analysis of RAPD patterns.

Limited genetic diversity was identified within MON-1 Portuguese *L. infantum* strains, but polymorphic profiles were produced by some primers (Fig. 21). Unfortunately most polymorphic bands were not strong enough to be reliably considered and were not scored.

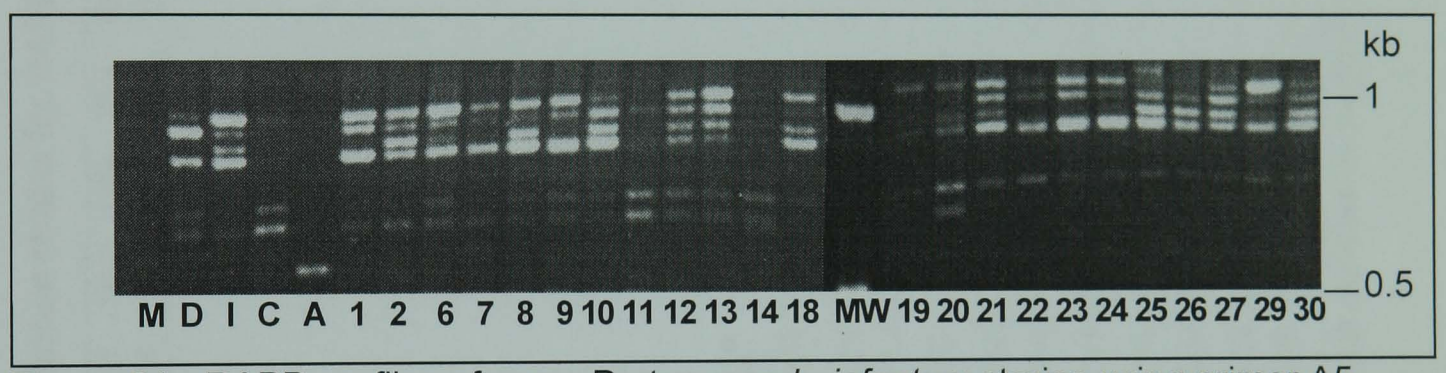


Figure 21 - RAPD profiles of some Portuguese *L. infantum* strains using primer A5.

Most Portuguese strains were more closely related than other *L. infantum* strains and MON 24 Portuguese strains, had Jaccard distances of between 0.24 and 0.60 (Table XXVIII), or 0.56 without MON 24 strains, as seen in Fig. 22. It was not possible to discriminate between different foci through UPGMA (CC = 0.99), single linkage (CC = 0.98) or neighbour joining trees, although three strains from the North clustered in the single linkage and neighbour joining trees. The phenograms were very assymetric, however, perhaps because the data were not sensitive enough to produce a phylogenetic structure or the sample was simply inadequate. Different methods generated different topologies within the Portuguese *L. infantum* strains and the neighbour-joining tree partially inverted the order of branching of these strains in the single linkage dendrogram (Fig. 22).

Table XVIII - Pairwise distance (Jaccard index) among Portuguese *L. infantum* strains for RAPD data.

	5-ASKH	L100	DD8	PP 75	IPT-1	L82	IMT 104	IMT 150	IMT 152	IMT 89	IMT 124	IMT 193	IMT 193	IMT 160	IMT 162	IMT 161	IMT 169	IMT 170	IMT 171	IMT 172	IMT 177	IMT 191	IMT 195
5-ASKH	**56**																						
L100	0.926	**56**																					
DD8	0.927	0.933	**74**																				
PP75	0.918	0.923	0.593	**76**																			
IPT-1	0.935	0.924	0.632	0.550	**70**																		
L82	0.927	0.915	0.625	0.556	0.370	**66**																	
IMT104	0.917	0.933	0.631	0.494	0.481	0.464	**75**																
IMT150	0.922	0.927	0.626	0.532	0.482	0.509	0.391	**82**															
IMT152	0.923	0.928	0.623	0.492	0.500	0.525	0.389	0.264	**84**														
IMT89	0.921	0.932	0.607	0.506	0.494	0.520	0.378	0.287	0.354	**81**													
IMT124	0.918	0.934	0.627	0.447	0.539	0.546	0.351	0.360	0.358	0.310	**77**												
IMT108	0.923	0.933	0.609	0.492	0.479	0.506	0.360	0.264	0.302	0.243	0.289	**84**											
IMT193	0.921	0.931	0.642	0.523	0.494	0.537	0.435	0.300	0.260	0.378	0.432	0.331	**88**										
IMT160	0.918	0.928	0.595	0.471	0.500	0.525	0.389	0.303	0.302	0.322	0.325	0.216	0.361	**84**									
IMT162	0.929	0.924	0.586	0.459	0.467	0.473	0.426	0.320	0.318	0.337	0.371	0.240	0.346	0.240	**85**								
IMT161	0.930	0.930	0.611	0.516	0.506	0.530	0.508	0.404	0.348	0.461	0.445	0.402	0.344	0.402	0.331	**87**							
IMT169	0.927	0.932	0.639	0.549	0.503	0.528	0.524	0.422	0.394	0.478	0.483	0.442	0.390	0.464	0.406	0.240	**82**						
IMT170	0.928	0.928	0.604	0.521	0.511	0.517	0.494	0.382	0.406	0.466	0.450	0.406	0.425	0.406	0.365	0.333	0.285	**83**					
IMT171	0.925	0.935	0.611	0.550	0.542	0.596	0.543	0.428	0.425	0.440	0.466	0.425	0.397	0.425	0.413	0.359	0.428	0.438	**87**				
IMTI72	0.925	0.930	0.645	0.537	0.547	0.603	0.548	0.450	0.469	0.462	0.445	0.447	0.484	0.447	0.457	0.454	0.450	0.437	0.322	**78**			
IMT177	0.930	0.925	0.624	0.508	0.476	0.503	0.500	0.413	0.457	0.450	0.431	0.384	0.452	0.411	0.339	0.367	0.358	0.371	0.464	0.391	**79**		
IMT191	0.917	0.933	0.618	0.521	0.511	0.517	0.429	0.285	0.320	0.397	0.343	0.320	0.376	0.320	0.335	0.333	0.354	0.303	0.415	0.437	0.371	**83**	
IMT195	0.917	0.932	0.626	0.532	0.522	0.564	0.506	0.397	0.367	0.411	0.415	0.367	0.390	0.394	0.380	0.378	0.369	0.408	0.404	0.426	0.413	0.322	**82**

The lower triangular matrix shows Jaccard distances, the diagonal the total number of fragments per strain.

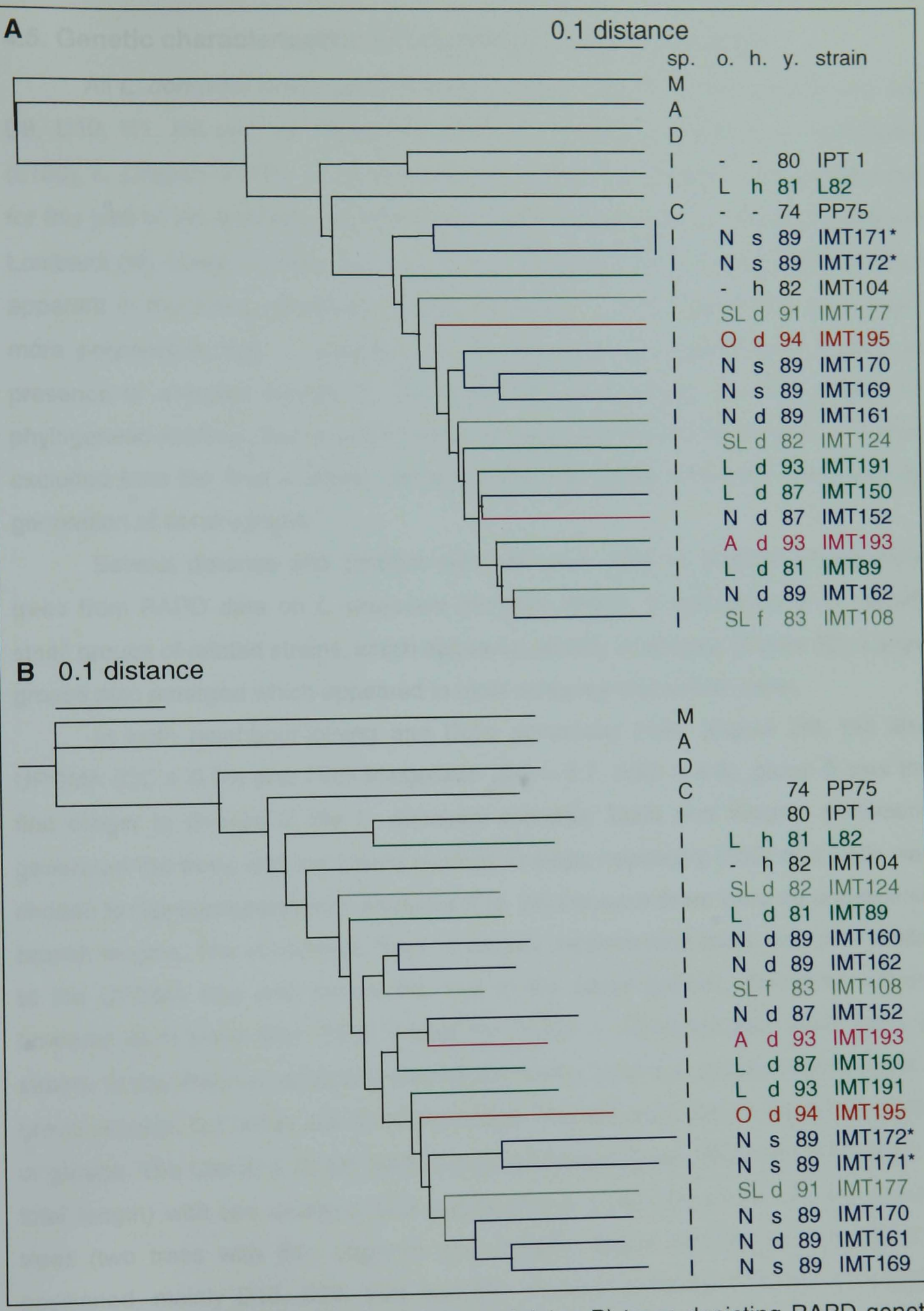


Figure 22 - Single linkage A) and neighbour-joining B) trees depicting RAPD genetic diversity within Portuguese *L. infantum*. Sp. - species; o. - geographical origin; h. - host; y. - year; A - *L. aethiopica* (L96), C - *L. chagasi* (PP75); D - *L. donovani* (DD8), I - *L. infantum* (IPT-1), M - *L. major* (5-ASKH); d - dog; f- fox; s - sandfly; h -human; A (purple) - Algarve, L (green) - Lisbon, N (blue) - North, O (red) - other; SL (dark green) - South Lisbon; * are zymodeme MON 24.

4.5. Genetic characterization of *Leishmania donovani* strains

All *L. donovani* strains were amplified by nine RAPD primers: A2, A4, A5, A6, D8, D10, H1, H4 and L2. Reference strains of *L. major* (5-ASKH), *L. aethiopica* (L100), *L. chagasi* (PP75), *L. infantum* (IPT1) were also included in the sample used for this part of the analysis, as well as the most divergent strains within *L. infantum*: Lombardi (I4), Strain A (I16), L53 (I17), IMT171 (I25) and IMT 172 (I26). As had been apparent in the first *L. donovani* complex analysis, *L. donovani* strains were much more polymorphic than *L. infantum* strains (Fig. 23). The bands were scored as presence or absence (Annex 2), individually for each primer and then pooled for phylogenetic analysis. Some strains did not amplify well with some primers, and were excluded from the final analysis. Jaccard distances (Table XIX) were calculated for generation of dendrograms.

Several distance and cladistic methods were used to produce phylogenetic trees from RAPD data on *L. donovani* complex strains. It was possible to identify small groups of related strains, which appeared reliably in all trees (Figure 22). Larger groups also emerged which appeared in most analyses (not shown here).

In both neighbour-joining and Dollo parsimony trees (Figure 24), but also UPGMA (CC = 0.94) and Fitch-Margoliash (SQ = 3.7; ASD = 4.8), group E was the first cluster to diverge in the *L. donovani* complex. Dollo and Wagner parsimony generated the trees with the lowest number of steps necessary (352) and Dollo was chosen to represent parsimony analyses (Fig. 24) because there were no undefinined branch lengths. The consensus, from 14 equally parsimonious trees, was very similar to the UPGMA tree and placed the root in the same position. Bootstrap values, however were lower than 65%, except for Indian *L. donovani* and some pairs of strains. In the Wagner parsimony method (24 trees), strains in group A did not form a group as such, but rather branched separately from the root, before any other strains or groups. The Camin & Sokal method produced trees (three trees, with 454 steps in total length) with two divisions and polymorphism parsimony produced the longest trees (two trees with 641 steps in total length). Some strains were not reliably positioned, mainly D16, D29, D24 and D5. In most methods (UPGMA and Dollo, Polymorphism, Wagner and Camin-Sokal parsimonies), D16 and D26 were placed with group A, whilst D26 was placed with group C by neighbour-joining. Strain D29 was placed with group C in UPGMA and group B in parsimony analysis. Although it was possible to identify clusters of genetically related strains in the distance matrix in Table XIX, definition of those groups was not very good.

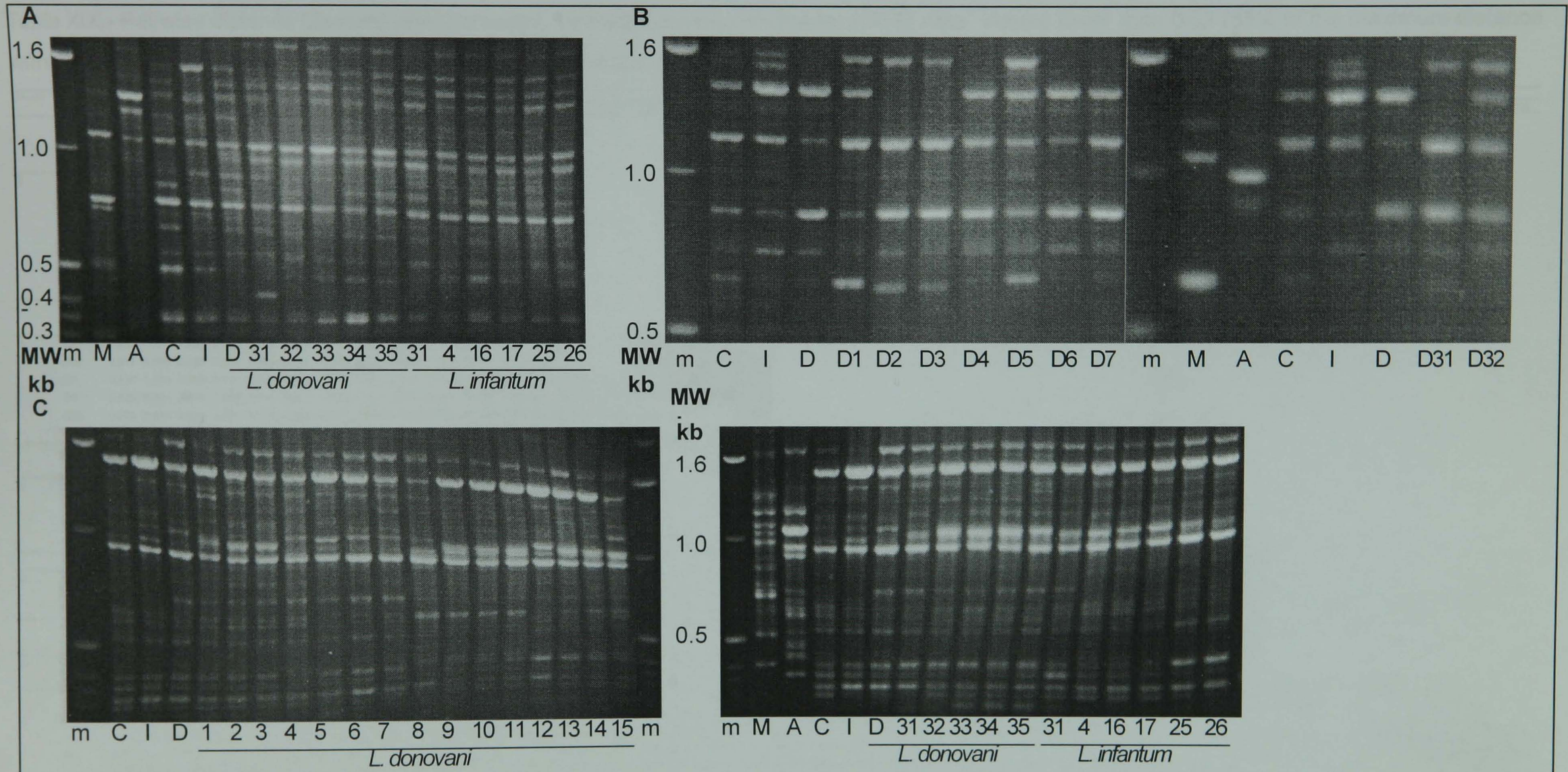


Figure 23 - RAPD band patterns within *L. donovani* with primers A) H4, B) D10 and C) L2. 'm' and 'MW' stand for molecular weight markers.

Table XIX - Pairwise distance (Jaccard index) among *L. donovani* complex strains for RAPD data. Values lower than 0.52 (65% of the maximum distance within the *L. donovani* complex) are shown in black on the upper half.

Group	Strain	M	A	D23	D17	D11	D15	D20	D12	D13	D10	D14	D26	D16	D21	D27	D2	D3	D25	D30	D24	D5	D	D4	D6	D7	D22	D32	D33	D34	D28	D29	D1	D19	C	I	I31	I4	I16	I17	I25	I26
	A	0.930	-																																						
	D23	0.961	0.906	-																																					
	D17	0.923	0.938	0.858	-																																				
E	D11	0.911	0.937	0.882	0.673	-																																			
	D15	0.913	0.926	0.864	0.643	0.571	-																																		
	D20	0.925	0.928	0.868	0.661	0.677	0.540	-																																	
	D12	0.921	0.944	0.885	0.707	0.622	0.623	0.526	-																																
	D13	0.905	0.945	0.878	0.677	0.615	0.589	0.484	0.374	-																															
	D10	0.918	0.921	0.849	0.678	0.522	0.461	0.456	0.520	0.479	-																														
	D14	0.907	0.938	0.871	0.670	0.544	0.484	0.438	0.422	0.365	0.330	-																													
	D26	0.918	0.930	0.894	0.740	0.770	0.688	0.661	0.737	0.713	0.713	0.707	-																												
	D16	0.918	0.921	0.884	0.678	0.714	0.566	0.614	0.680	0.629	0.632	0.624	0.609	-																											
A	D21	0.931	0.935	0.898	0.714	0.784	0.679	0.600	0.648	0.616	0.707	0.659	0.644	0.594	-																											
	D27	0.917	0.929	0.893	0.693	0.685	0.583	0.653	0.694	0.644	0.688	0.638	0.572	0.572	0.583	-																										
	D2	0.930	0.934	0.899	0.720	0.733	0.624	0.662	0.701	0.655	0.695	0.649	0.588	0.532	0.510	0.404	-																									
	D3	0.928	0.931	0.886	0.707	0.758	0.653	0.624	0.688	0.638	0.682	0.632	0.593	0.566	0.479	0.452	0.400	-																								
	D25	0.923	0.934	0.901	0.726	0.756	0.632	0.670	0.725	0.701	0.701	0.695	0.480	0.514	0.555	0.464	0.416	0.495	-																							
	D30	0.919	0.931	0.896	0.727	0.740	0.604	0.647	0.688	0.638	0.682	0.655	0.537	0.469	0.515	0.408	0.400	0.447	0.374	-																						
F	D24	0.926	0.929	0.913	0.734	0.798	0.754	0.694	0.694	0.707	0.757	0.719	0.542	0.647	0.659	0.614	0.577	0.632	0.561	0.609	-																				
	D5	0.930	0.925	0.899	0.720	0.784	0.725	0.662	0.661	0.675	0.730	0.689	0.588	0.612	0.624	0.577	0.567	0.598	0.551	0.598	0.404	-																		
B	D	0.938	0.916	0.909	0.772	0.768	0.707	0.641	0.638	0.655	0.676	0.627	0.612	0.635	0.647	0.649	0.592	0.621	0.602	0.643	0.577	0.509	-																	
	D4	0.937	0.914	0.898	0.751	0.745	0.681	0.655	0.653	0.647	0.670	0.618	0.577	0.627	0.661	0.618	0.556	0.612	0.592	0.612	0.566	0.556	0.319	-																
	D6	0.936	0.903	0.906	0.763	0.775	0.694	0.647	0.667	0.681	0.682	0.632	0.618	0.566	0.653	0.632	0.572	0.603	0.582	0.603	0.555	0.514	0.350	0.464	-															
	D7	0.936	0.913	0.886	0.727	0.740	0.674	0.647	0.644	0.638	0.662	0.609	0.593	0.537	0.653	0.609	0.544	0.577	0.582	0.577	0.524	0.514	0.400	0.378	0.292	-														
	D22	0.922	0.935	0.909	0.692	0.748	0.679	0.540	0.595	0.559	0.644	0.583	0.644	0.620	0.606	0.701	0.668	0.674	0.655	0.629	0.583	0.599	0.647	0.638	0.629	0.629	-													
C	D32	0.917	0.928	0.894	0.701	0.714	0.627	0.544	0.618	0.561	0.592	0.556	0.592	0.476	0.649	0.607	0.597	0.602	0.606	0.577	0.630	0.572	0.546	0.535	0.577	0.551	0.577	-												
	D33	0.920	0.924	0.898	0.693	0.745	0.661	0.555	0.604	0.572	0.627	0.593	0.627	0.549	0.614	0.618	0.607	0.612	0.616	0.588	0.618	0.527	0.556	0.598	0.561	0.588	0.529	0.392	-											
	D34	0.918	0.921	0.894	0.678	0.770	0.688	0.588	0.659	0.629	0.655	0.647	0.583	0.555	0.620	0.668	0.657	0.618	0.598	0.618	0.599	0.532	0.588	0.627	0.593	0.618	0.535	0.439	0.292	-										
	D28	0.928	0.940	0.886	0.639	0.740	0.674	0.599	0.667	0.614	0.662	0.632	0.593	0.618	0.604	0.609	0.572	0.603	0.582	0.577	0.555	0.514	0.572	0.561	0.603	0.577	0.515	0.551	0.464	0.505	-									
	D29	0.910	0.931	0.886	0.727	0.700	0.629	0.599	0.620	0.560	0.618	0.583	0.537	0.505	0.604	0.555	0.514	0.603	0.527	0.447	0.583	0.572	0.572	0.500	0.577	0.519	0.548	0.491	0.532	0.593	0.485	-								
G	D1	0.923	0.934	0.910	0.687	0.756	0.695	0.627	0.647	0.618	0.683	0.635	0.572	0.598	0.632	0.612	0.577	0.630	0.562	0.607	0.500	0.456	0.522	0.509	0.582	0.556	0.524	0.500	0.509	0.544	0.460	0.495	-							
	D19	0.914	0.927	0.910	0.700	0.714	0.665	0.553	0.546	0.505	0.629	0.566	0.629	0.604	0.616	0.620	0.609	0.614	0.641	0.588	0.594	0.583	0.583	0.572	0.614	0.588	0.489	0.500	0.433	0.515	0.529	0.495	0.429	-						
D	C	0.927	0.93	0.933	0.775	0.786	0.725	0.701	0.701	0.694	0.713	0.688	0.632	0.655	0.688	0.725	0.695	0.701	0.643	0.662	0.624	0.612	0.588	0.603	0.618	0.618	0.594	0.638	0.649	0.609	0.618	0.593	0.572	0.629	-						
	I	0.929	0.932	0.926	0.800	0.797	0.729	0.707	0.707	0.701	0.733	0.696	0.702	0.685	0.729	0.696	0.702	0.723	0.691	0.691	0.696	0.630	0.609	0.642	0.655	0.691	0.658	0.632	0.622	0.648	0.655	0.635	0.596	0.624	0.500	-				
	I31	0.934	0.929	0.913	0.752	0.765	0.680	0.629	0.714	0.688	0.688	0.681	0.624	0.572	0.659	0.661	0.649	0.675	0.635	0.632	0.661	0.603	0.627	0.641	0.632	0.655	0.583	0.556	0.505	0.510	0.555	0.555	0.588	0.594	0.542	0.572	-			
	I4	0.937	0.933	0.917	0.733	0.763	0.661	0.632	0.694	0.688	0.649	0.641	0.627	0.627	0.701	0.682	0.671	0.676	0.638	0.657	0.662	0.630	0.630	0.621	0.635	0.635	0.588	0.587	0.572	0.577	0.561	0.561	0.539	0.572	0.519	0.577	0.429	-		
	I16	0.936	0.931	0.925	0.745	0.791	0.694	0.647	0.725	0.701	0.682	0.675	0.593	0.593	0.653	0.675	0.643	0.649	0.582	0.603	0.632	0.598	0.598	0.612	0.577	0.603	0.548	0.577	0.532	0.505	0.519	0.549	0.556	0.614	0.469	0.592	0.357	0.323	-	
	I17	0.936	0.940	0.916	0.727	0.791	0.713	0.647	0.707	0.681	0.701	0.695	0.566	0.593	0.629	0.655	0.621	0.627	0.556	0.577	0.609	0.572	0.598	0.612	0.603	0.603	0.548	0.577	0.532	0.505	0.485	0.519	0.527	0.588	0.469	0.592	0.357	0.378	0.209	-
	I25	0.939	0.943	0.901	0.744	0.756	0.675	0.649	0.668	0.662	0.664	0.657	0.643	0.572	0.655	0.657	0.645	0.671	0.611	0.607	0.676	0.577	0.602	0.638	0.630	0.630	0.609	0.557	0.509	0.544	0.556	0.527	0.535	0.566	0.544	0.548	0.378	0.396	0.420	0.374	-
	I26	0.940	0.944	0.902	0.732	0.761	0.682	0.657	0.675	0.670	0.671	0.664	0.630	0.582	0.662	0.643	0.632	0.658	0.597	0.616	0.664	0.562	0.587	0.624	0.616	0.616	0.618	0.541	0.491	0.527	0.539	0.539	0.518	0.577	0.556	0.558	0.400	0.370	0.396	0.346	0.143

D34 = D35. A - *L. aethiopica*; M - *L. major*

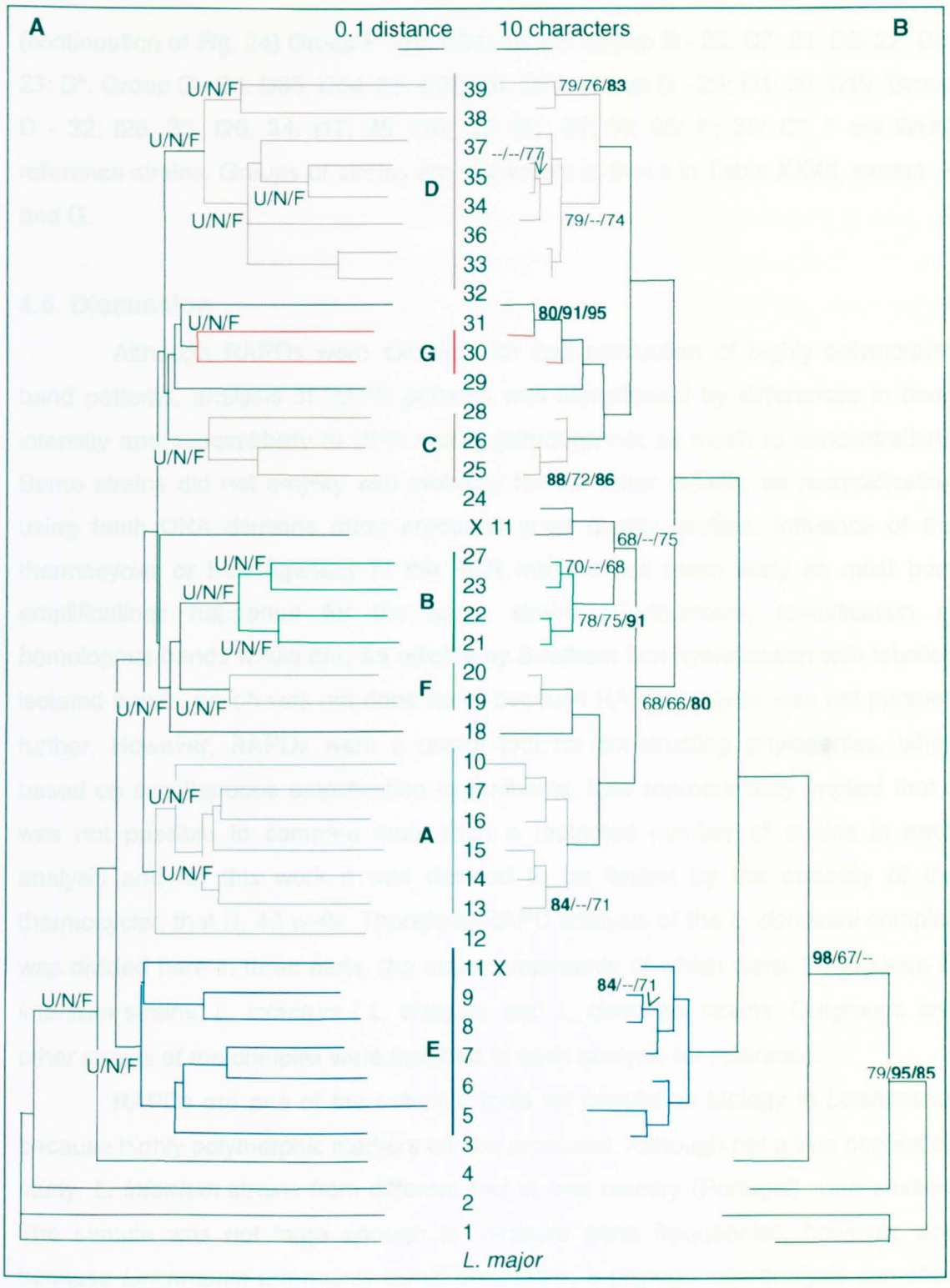


Figure 24 - Phylogenetic trees of the *L. donovani* complex based on RAPD data. A) Neighbour-joining and B) Consensus Dollo parsimony tree. Bootstrap values higher than 65% are shown, from 100 replicate data sets, for Dollo/ Wagner/ polymorphism parsimony. Branches present in both UPGMA / neighbour-joining / Fitch Margoliash trees are indicated by U/N/F. Ungrouped - 1: D23; 2: D17; 10: D26; 11: D16; 27: D29; 28: D28; 31: D22. Group E - 3: D11; 4: D15; 5: D10; 6: D14; 7: D12; 8: D13; 9: D20. Group A - 12: D25; 13: D30; 14: D27; 15: D2; 16: D3; 17: D21. (continues next page)

(continuation of Fig. 24) Group F - 18: D24; 19: D5. Group B - 20: D7; 21: D6; 22: D4; 23: D*. Group C - 24: D35, D34; 25: D33; 26: D32. Group G - 29: D1; 30: D19. Group D - 32: I26; 33: I25; 34: I17; 35: I16; 36: I31; 37: I4; 38: I*; 39: C*. * are WHO reference strains. Groups of strains are equivalent to those in Table XXXII, except F and G.

4.5. Discussion

Although RAPDs were excellent for fast production of highly polymorphic band patterns, analysis of RAPD patterns was complicated by differences in band intensity and susceptibility to DNA quality (although not so much to concentration). Some strains did not amplify well probably for the latter reason, as reamplification using fresh DNA dilutions often produced good quality profiles. Influence of the thermocycler or heterogeneity of the PCR mix did not seem likely as most poor amplifications happened for the same strains. Furthermore, identification of homologous bands would only be reliable by Southern blot hybridization with labelled isolated bands, which was not done here, because RAPD analysis was not pursued further. However, RAPDs were a useful tool for constructing phylogenies, when based on simultaneous amplification experiments. Low reproducibility implied that it was not possible to compare more than a restricted number of strains in each analysis and for this work it was decided to be limited by the capacity of the thermocycler, that is, 46 wells. Therefore, RAPD analysis of the *L. donovani* complex was divided here in three parts, the main components of which were: Portuguese *L. infantum* strains; *L. infantum / L. chagasi,* and *L. donovani* strains. Outgroups and other strains of the complex were included in each analysis for reference.

RAPDs are one of the potential tools for population biology in *Leishmania*, because highly polymorphic markers can be produced. Although not a true population study, *L. infantum* strains from different foci of one country (Portugal) were studied. The sample was not large enough to measure gene frequencies, however, and because *Leishmania* are mainly clonal organisms, a phylogenetic analysis can shed some light on population structure, or on evolutionary independence of different populations. By clustering methods using Jaccard distances it was not possible to individualize any one population, meaning that no population was sufficiently differentiated from the others or that this RAPD analysis was not sufficiently sensitive. The structures of the dendrograms and cladograms were very assymetric, however, and it is possible that not enough information was obtained to scrutinize the sample with high sensitivity or, conversely, that the method was not adequate, simply because the genetic diversity level was indeed very low in the sample. The results,

however, indicate that Portuguese foci are probably not isolated and that control in one area may be hindered by immigration of isolates from other foci. If populations are indeed separated, the results suggest that those populations must have only recently separated and have not yet fully differentiated. The *Leishmania* genetic pool in Portugal or in each geographical location taken as a population seems to have, at least at present, an endemic rather than an epidemic structure, meaning that diverse genotypes are present rather than only a few different genotypes favoured in epidemic outbreaks.

In an extended analysis of genetic diversity by RAPD within *L. infantum* and *L. chagasi*, it was not possible to discriminate the two species, although diversity was found. The close relationship between *L. infantum* and *L. chagasi* contrasted with a higher level of genetic diversity of *L. donovani*, the tested strains of which were easily separated from the *L. infantum / L. chagasi* group and from each other. The impossibility of distinguishing here between *L. infantum* and *L. chagasi* strains, with similar intraspecific variability within each, gave strong support to the hypothesis that these species are synonymous (Momen *et al.,* 1987). Furthermore, *L. infantum / L. chagasi* formed a single branch in RAPD dendrograms suggesting that *L. infantum / L. chagasi* is a well individualized, monophyletic group.

A substantial amount of diversity was observed within the *L. donovani* strains included on the above analysis. On the UPGMA dendrogram, the *L. donovani* complex was rooted within *L. donovani* strains, with *L. infantum* depicted as a recent branch of *L. donovani*, suggestive of *L. donovani* as a paraphyletic group, contrasting with earlier findings (Moreno *et al.,* 1984; Thomaz Soccol *et al.,* 1993), but supporting others (Le Blancq *et al.,* 1986; Mebrahtu *et al.,* 1992).

The need for an extended analysis of *L. donovani* resulted in the analysis of a total of 30 *L. donovani* strains and six *L. infantum / L. chagasi* strains. The study of the *L. donovani* complex evidenced division of the taxa into seven groups of which only the Indian cluster was robust in parsimony analyses. Most distance based methods agreed with each other on group formation, especially groups D, C, B-F, A and E. Group C-G was not reliably identified, perhaps because of strain D28 (Gebre 1), which was also closer to D16 (Ndandu 4A). This latter could appear closer to group A. Groups B, F and A seemed to be related and boostrap values were higher than 65% for the lineage containing these three groups.

The overall low bootstrap values suggest that RAPDs may be too sensitive for phylogenenetic analyses. Alternatively, it may be that RAPD data does not produce reliable synapomorphisms to be amenable for parsimony analyses and some of the indefinitions on the RAPD tree might be due to recombination events, which

would have to be studied extensively. However, RAPD seemed very useful for identification of genetic diversity between closely related strains. In the history of the *L. donovani* complex as shown by RAPD, group E was seemingly the first to diverge from the initial *L. donovani* genetic pool, with four groups emerging later in a probable radiation event, or a five clade radiation event may also have occurred. The different RAPD analyses suggested that emergence of *L. infantum* / *L. chagasi* was a recent episode in the life of the complex, that *L. chagasi* is not distinct from *L. infantum* and also indicated paraphyly of *L. donovani*.

5. Comparison of Intergenic Regions: PCR - RFLP

5.1. Introduction

5.1.1. DNA restriction

Most bacterial strains have enzymes capable of recognising specific DNA sequences and either methylating bases (adenine or cytosine residues) in a (usually recently synthesised) non-methylated DNA strand or of cutting DNA non-methylated in both strands. This means that foreign DNA which does not have the DNA modifications specific to the strain of bacterium is susceptible to attack by restriction enzymes of that bacterium. These restriction endonucleases were discovered when cross-infection with bacteriophages failed to succeed, and it is thought that such enzymes may play a role in protecting against bacteriophage infections or against any other foreign DNA, for example blocking conjugation with incompatible strains.

Most restriction enzymes (type II) recognise and cut, or methylate, within or near a specific short DNA sequence (usually 4 or 6bp), thus functioning as separate endonucleases and methylases. As a rule, endonucleases recognise palindromic sequences, but a few recognise uneven, nonpalindromic binding sites, or allow for one or more non specific nucleotides. Enzymes with similar recognition sequences may be sensitive to different methylation patterns and may cut in different positions: these features may be useful for different molecular biology applications.

The restriction pattern of genomic DNA or of a DNA fragment, clearly is sequence dependent, thus, appearance or disappearance of restriction sites and shortening or elongation of restriction fragments enables indirect comparison of DNA sequences. Different restriction patterns are called restriction fragment length polymorphism (RFLP) and are used for several purposes, including strain typing, DNA fingerprinting, parent determination, etc.

One way of producing restriction maps is to cut whole genomic DNA, separate the fragments by electrophoresis and transfer the DNA onto a membrane (Southern blot) which is then probed with appropriately labelled DNA probes to identify the desired fragments. This method provided some of the first molecular data on genetic organisation and genetic diversity of *Leishmania* (Beverley *et al.*, 1984; Macedo *et al.*, 1992; Mendoza Leon *et al.*, 1995; Oskam *et al.*, 1998; Ramirez and Guevara, 1987; van Eys *et al.*, 1991; van Eys *et al.*, 1989).

With the invention of PCR, DNA fragments from a specific region can be obtained in large quantities and restriction fragments produced from PCR products can be readily analysed by gel electrophoresis. Series of restrictions with different

enzymes allow comparisons of DNA sequence from different individuals in what became known as PCR-RFLP. This is a very powerful technique for comparison of DNA regions which are difficult or too long to sequence or when there is a large number of samples to be compared. PCR-RFLP requires knowledge of the flanking sequences for primer design, although not of the amplified region.

The converse technique was named amplified restriction fragment polymorphism (AFLP). Whole DNA is restricted with two enzymes, one a six and the other a four base cutter, and adapters are ligated to the ends of the fragments. Fragments are amplified using one labelled primer for the six base cutter end and one unlabelled primer for the four base cutter end. Both primers have sequences complementary to the adapters. Double digest fragments are detected by autoradiography or another detection system, after separation in a gel. The fragments generated by the six base cutter are most likely too long to be productively amplified by PCR, while fragments produced by the four base cutter will not be detected because the primer is not labelled. This method does not require the large amounts of DNA needed for classical RFLP, nor the knowledge of sequence demanded by PCR-RFLP, but still requires detection by labelled markers.

Restriction analysis is very robust and reproducible, but it can only identify sequence differences at the recognition site or differences in fragment length. To differentiate further between similar sized fragments other techniques must be employed, such as denaturing gradient gel electrophoresis (DGGE), single stranded conformational polymorphisms (SSCP) and DNA sequencing (see Introduction).

5.1.2. Ribosomal internal transcribed spacer

Leishmania rRNA genes are arranged in tandem repeat units, separated by a non-transcribed spacer (NTS) region (Cupolillo *et al.*, 1995). Each repeat has coding regions for the small sub-unit (SSU) or 18S rRNA and the large sub-unit (LSU) or 28S rRNA, which are separated by an internal transcribed spacer (ITS). The ITS region is divided into ITS1 and ITS2 by the coding region for the small 5.8S. *T. brucei* (Rudenko, 1999) and *C. fasciculata* (Spencer *et al.*, 1987) were found to have about six copies of LSU genes separated by spacers. Both NTS and ITS, as well as the LSU spacers, are postulated to suffer less selective pressure against mutations than the highly conserved coding regions and were found to be more polymorphic than the coding regions. The ITS is particularly interesting as it is flanked by highly conserved segments for which PCR primers can be designed and it is relatively small (about 1kb in *Leishmania*). Study of the ITS located between the SSU and the first LSU gene was reported (Cupolillo *et al.*, 1995) for several American *Leishmania* species with

good correlation with isoenzyme and mini-exon data.

5.1.3. Mini-exon

The mini-exon gene codes for the spliced leader of all kinetoplastid mRNA (Nilsen, 1994). It is present in 100-200 tandem repeat gene units (Fernandes *et al.*, 1994), each with a transcribed region and a non transcribed spacer. The transcribed region codes for a highly conserved 39 nucleotide exon, the RNA transcript of which is trans-spliced to the 5'-end of all nuclear mRNA (Fernandes *et al.*, 1994), and a less conserved intron of between 55 to 101 nucleotides. The non-transcribed spacer is highly variable in sequence (ranging from 250 to 1350 nucleotides), and is thus useful to discriminate between closely related species of kinetoplastids. In the *L. donovani* complex the amplification of the mini-exon repeat unit generated products from 422 to 467bp (Fernandes *et al.*, 1994). This region has been used successfully in PCR diagnosis (Hassan *et al.*, 1993; Katakura *et al.*, 1998), and for identification of *Leishmania* complexes (Harris *et al.*, 1998; Ramos *et al.*, 1996). The mini-exon is a potentially good target for typing, since it shows variability within the *L. donovani* complex (Fernandes *et al.*, 1994) but there are no intra-specific studies as yet.

5.1.4. Intergenic regions of *msp* genes

The *Leishmania* major surface protease (*msp)* genes are multicopy in all studied *Leishmania* species. Gene organization has not been described, except for *L. chagasi* (Roberts *et al.*, 1993) and *L. guyanensis* (Steinkraus *et al.*, 1993) and to a lesser degree for *L. mexicana* (Medina Acosta *et al.*, 1993), although different gene families could be identified (Maingon *et al.*, 1990; Victoir *et al.*, 1995; Voth *et al.*, 1998). In *L. chagasi*, three gene families were described, which initially were said to be expressed preferentially in logarithmic or stationary growth phases, or to be constitutively expressed (Ramamoorthy *et al.*, 1992). Gene families were accordingly named *mspL*, *mspS* and *mspC*, although it was later found that all genes were expressed in all stages of development (Roberts *et al.*, 1993). *MspC* was identified as a single copy gene, *mspL* were present in an array of identical genes and *mspS* were found to have at least 4 subgroups. Each gene family can be distinguished by its non-coding region and the 3' half of the coding region. *Msp* or gp63 genes were found to be located on chromosome 10. In *L. chagasi*, there is an array of an indeterminate number of *mspS*, followed by an array of an unknown number of *mspL*. Next is *mspC*, followed by a single copy of *mspS4*. Sequences for *L. infantum* gp63 were obtained recently (Gonzalez Aseguinolaza *et al.*, 1997; Morales *et al.*, 1997).

5.2. Ribosomal internal transcribed spacer

5.2.1. PCR amplification

The sizes of the *L. donovani* complex PCR products were, in agarose gel electrophoresis, approximately, 1kb (Fig. 25). The alternative primers ITS1F and ITS2R amplified a product shorter by 34 bp (sITS), which was comparable to the full ITS product in terms of yield and the appearance of a single band using similar amplification conditions. A fragment of approximately 1.1kb was amplified from *L. major*, one fragment of 1.05kb from *L. aethiopica* and two fragments of 1.05kb and 0.9kb from *L. tropica*.

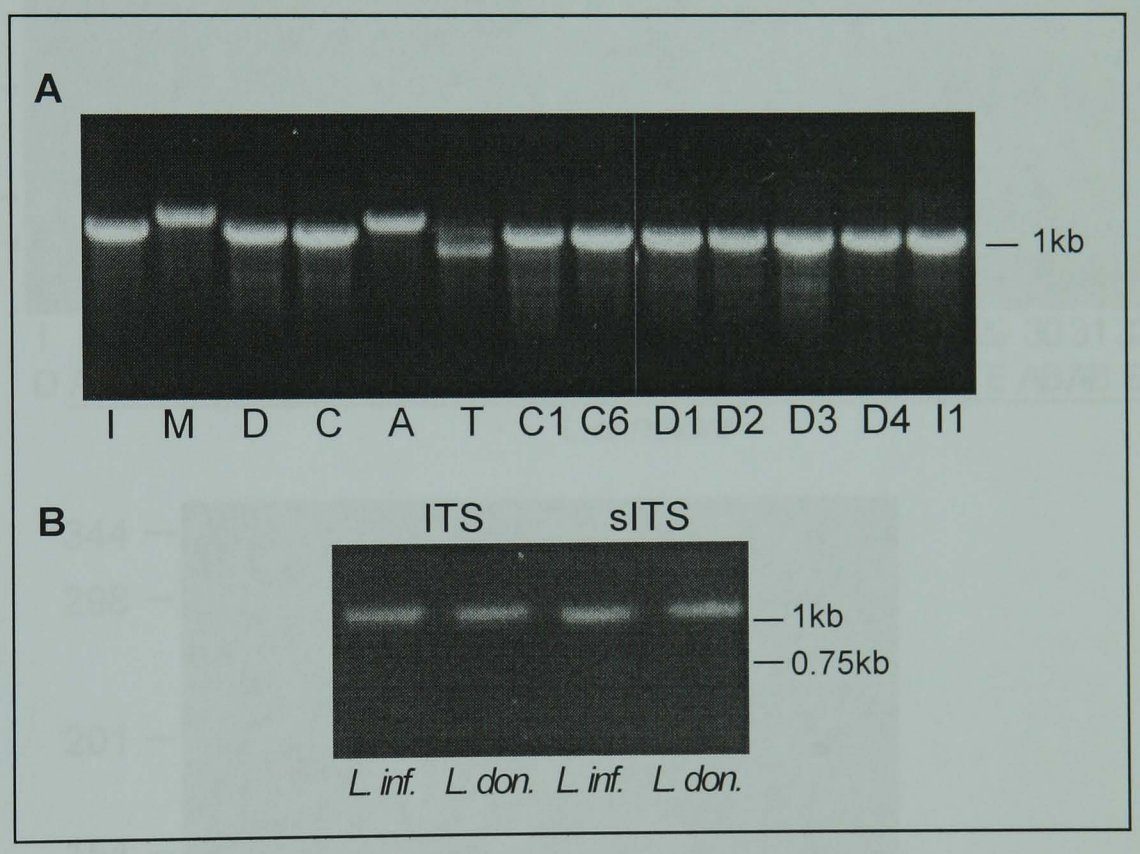


Figure 25 - Amplification products from A) ITS within the *L. donovani* complex and other OW *Leishmania* reference strains, and B) sITS compared with ITS. Reference strains: A - *L. aethiopica*; C - *L. chagasi*; I - *L. infantum*; D - *L. donovani*; M - *L. major*; T - *L. tropica*. For strain codes, see Tables II to V.

5.2.2. Restriction analysis of the internal transcribed spacer in the *Leishmania donovani* complex

Only small size differences were observed in the ITS PCR-RFLP profiles within the *L. donovani* complex (Fig. 26). The outgroup strains, by contrast, were highly divergent. Enzymes *Bst*UI, *Eco*RI, *Hae*III, *Mse*I and *Taq*I were the most polymorphic. The full RFLP data are shown in Table XL (Annex 2). Restriction fragments for most strains added up to the expected size with *Alu*I, *Hae*III and *Mse*I (Table XX), allowing for small errors in determination of band size, although with other enzymes the total

fragment size was much larger. Strains D9, D10, D11, D13, D20, D26, D29 and D32, gave fragments that added up to approximately twice the expected size with enzymes *Eco*RI, *Hae*III, *Sph*I and *Taq*I.

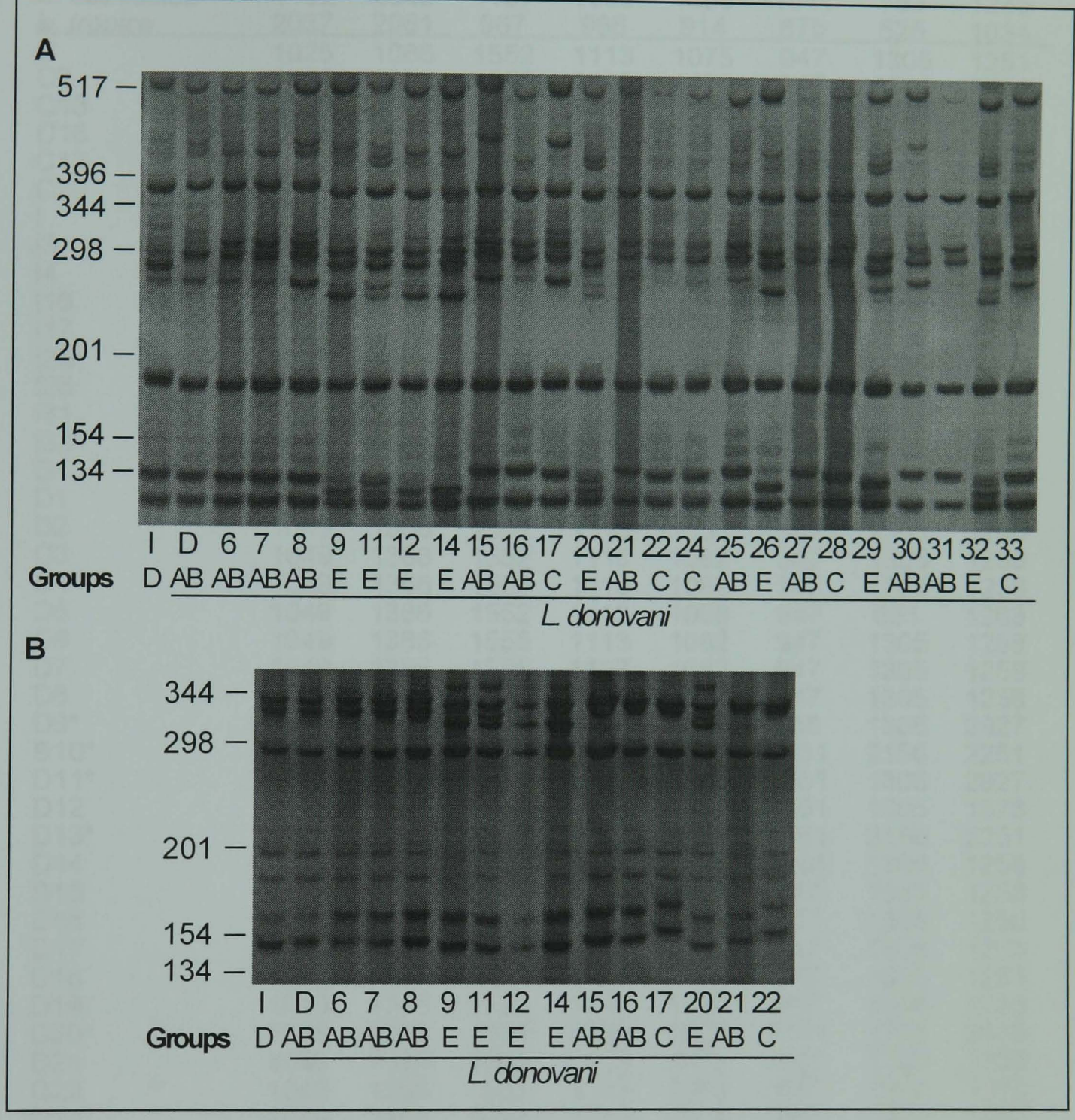


Figure 26 - Restriction profiles of ITS obtained with A) *Bst*UI and B) *Taq*I within the *L. donovani* complex and outgroup strains. Strain codes are as in Tables IV and V. Genetic groups are shown for each profile and are as shown in Table XXXII (insert).

Table XX - Total fragment lengths (bp) scored per strain and per enzyme for ITS RFLP in the *L. donovani* complex.

Strain \ Enzyme	*Alu*I	*Bst*UI	*Cfo*I	*Eco*RI	*Hae*III	*Mse*I	*Sph*I	*Taq*I
L. maior	1161	1107	546	1153	1243	1137	1081	1285
L. aethiopica	1102	1545	1137	1099	1099	1044	753	1235
L. tropica	2037	2061	967	966	914	875	525	1034
C	1025	1086	1552	1113	1075	947	1305	1251
C2	1025	1086	1552	1113	1075	947	1305	1251
C13	1025	1086	1552	1113	1075	947	1305	1251
C16	1025	1086	1552	1113	1075	947	1305	1251
C17	1025	1086	1552	1113	1075	947	1305	1251
C20	1025	1366	1552	1113	1075	947	1305	1251
I	1025	1086	1552	1113	1075	947	1305	1251
I3	1025	1086	1552	1113	1075	947	1305	1251
I4	1025	1086	1552	1113	1075	947	1305	1251
I16	1025	1086	1552	1113	1075	947	1305	1251
I17	1025	1086	1552	1113	1075	947	1305	1251
I25	1025	1086	1552	1113	1075	947	1305	1251
I26	1025	1086	1552	1113	1075	947	1305	1251
I31	1025	1366	1552	1113	1075	947	1305	1251
I33	1025	1366	1552	1113	1075	947	1305	1251
D	1049	1366	1555	1113	1082	947	1305	1258
D1	1049	1366	1552	1113	1082	947	1305	1258
D2	1049	1366	1552	1113	1082	947	1305	1258
D3	1049	1366	1552	1113	1082	947	1305	1258
D4	1049	1366	1555	1113	1082	947	1305	1258
D5	1049	1386	1552	1113	1088	947	631	1263
D6	1049	1386	1555	1113	1082	947	1305	1258
D7	1049	1386	1555	1113	1082	947	1305	1258
D8	1049	1386	1555	1113	1082	947	1305	1258
D9*	1049	1357	1552	1098	2087	938	1305	2927
D10*	1049	1366	1552	2269	2087	1161	2156	2251
D11*	1049	1489	1552	2269	1695	1161	1305	2927
D12	1049	1357	1552	1098	1448	1161	1305	1573
D13*	1049	1366	1552	2269	2087	1161	2156	2251
D14	1049	1357	1552	1098	1065	1161	1305	1258
D15	1049	1388	1552	1113	1082	947	1305	1258
D16	1049	1388	1552	1113	1082	947	1305	1258
D17	1049	1388	1552	1113	1088	947	1305	1263
D18	1049	1386	1552	1113	1088	947	1305	1263
D19	1049	1386	1552	1113	1088	947	1305	1263
D20*	1049	1489	1552	2269	1695	1161	1305	2920
D21	1049	1388	1552	1113	650	947	1305	1258
D22	1049	1386	1552	1113	1088	947	1305	1263
D24	1049	1386	1552	1113	1088	947	1305	1263
D25	1049	1388	1552	1113	1082	947	1305	1258
D26*	1049	1364	1552	1098	1065	938	1305	2920
D27	1049	1388	1552	1113	1082	947	1305	1258
D28	1049	1388	1552	1113	1088	947	1305	1263
D29*	1049	1366	1552	2269	1695	1161	1305	2920
D30	1049	1388	1552	1113	1082	947	1305	1258
D31	1049	1388	1552	1113	1082	947	1305	1258
D32*	1049	1199	1552	2269	1695	1161	1305	2920
D33	1049	1368	1552	1113	988	947	1305	1246
D34	1025	1386	1552	1113	1088	947	1305	1263
D35	1025	1386	1552	1113	1088	947	1305	1263

Fragment size was determined by comparison with molecular markers or extrapolation. Accumulation of small errors may cause some variation in total size differences, whilst heterogeneous targets may produce artificial larger sizes (*).

Table XXI - Pairwise distance (Jaccard index) among *L. donovani* complex strains for ITS RFLP data. Pairs of strains with distance values lower than 0.484 (65% of the maximum distance within the *L. donovani* complex) are shown in black in the upper half.

Group					E									D		C				AB				
		L. m.	*L. a.*	*L. t.*	D26	D10 D13	D11	D20	D29	D32	D9	D12	D14	li/lc	I33 I31 C20	D33	D34 D35	D17 D28	D5,D18 D19 D22 D24	D D4	D1 D2 D3	D6 D7 D8	D15,D16 D25, D27,D30 D31	D21
	L. major (L. m.)	-																						
	L. aethiopica (L. a.)	0.964	-																					
	L. tropica (L. t.)	0.973	0.905	-																				
E	D26	0.967	0.985	0.952	-																			
	D10, D13	0.970	0.987	0.948	0.505	-																		
	D11	0.970	0.987	0.949	0.520	0.429	-																	
	D20	0.970	0.987	0.949	0.442	0.326	0.292	-																
	D29	0.970	0.987	0.948	0.422	0.295	0.326	0.149	-															
	D32	0.970	0.987	0.956	0.466	0.357	0.326	0.149	0.211	-														
	D9	0.968	0.986	0.954	0.452	0.515	0.361	0.456	0.479	0.479	-													
	D12	0.966	0.985	0.951	0.506	0.559	0.422	0.505	0.526	0.526	0.378	-												
	D14	0.965	0.985	0.950	0.512	0.595	0.471	0.546	0.565	0.565	0.436	0.232	-											
D	*L. infantum / L. chagasi* (li/lc)	0.973	0.984	0.948	0.591	0.655	0.686	0.639	0.629	0.655	0.677	0.640	0.617	-										
	I33, I31, C20	0.964	0.984	0.939	0.571	0.639	0.672	0.623	0.612	0.639	0.661	0.622	0.598	0.171	-									
C	D33	0.953	0.984	0.948	0.650	0.721	0.686	0.707	0.700	0.721	0.652	0.610	0.584	0.555	0.531	-								
	D34, D35	0.974	0.984	0.949	0.707	0.745	0.733	0.733	0.727	0.745	0.728	0.700	0.683	0.487	0.506	0.487	-							
	D17, D28	0.964	0.984	0.958	0.684	0.745	0.733	0.733	0.727	0.745	0.707	0.676	0.657	0.570	0.584	0.373	0.333	-						
	D5, D18, D19, D22, D24	0.974	0.984	0.949	0.684	0.727	0.714	0.714	0.707	0.727	0.707	0.676	0.657	0.531	0.548	0.435	0.239	0.239	-					
AB	D, D4	0.964	0.976	0.939	0.684	0.727	0.672	0.714	0.707	0.727	0.661	0.622	0.598	0.570	0.548	0.487	0.584	0.584	0.548	-				
	D1, D2, D3	0.964	0.984	0.939	0.659	0.707	0.648	0.693	0.686	0.707	0.636	0.591	0.563	0.531	0.506	0.435	0.548	0.548	0.506	0.239	-			
	D6, D7, D8	0.974	0.976	0.949	0.707	0.745	0.693	0.733	0.727	0.745	0.685	0.650	0.629	0.570	0.584	0.531	0.548	0.548	0.506	0.239	0.333	-		
	D15, D16, D25, D27, D30, D31	0.964	0.984	0.958	0.684	0.745	0.693	0.733	0.727	0.745	0.661	0.622	0.598	0.570	0.584	0.435	0.548	0.459	0.506	0.403	0.333	0.333	-	
	D21	0.962	0.984	0.956	0.657	0.728	0.670	0.714	0.707	0.728	0.632	0.584	0.555	0.562	0.577	0.414	0.538	0.441	0.493	0.493	0.441	0.441	0.297	-

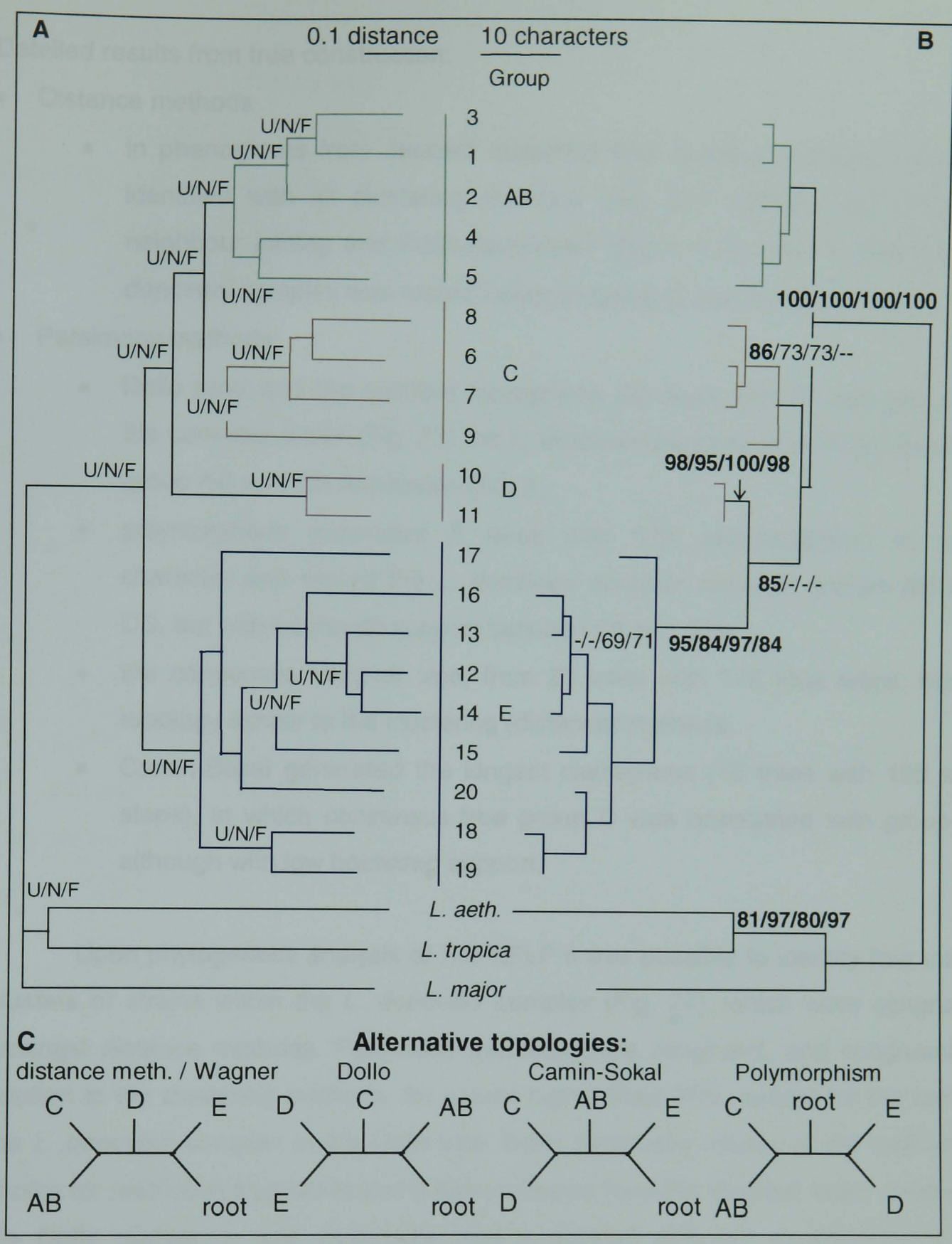


Figure 27 - Phylogenetic trees of the *L. donovani* complex produced from ITS RFLP data with alternative topologies (C). A) Neighbour-joining tree from Jaccard distances. Branches present in all UPGMA, neighbour-joining and Fitch-Margoliash trees (U/N/F) are indicated. B) Dollo parsimony tree. Bootstrap support higher than 70% (higher than 80% in bold) is shown for Dollo / Wagner / polymorphism / Camin-Sokal parsimony. Group AB: 1: D*, D4; 2: D1, D2, D3; 3 - D6, D7, D8; 4: D15, D16, D25, D27, D30, D31; 5: D21. Group C: 6: D5, D18, D19, D22, D24; 7: D17, D28; 8: D34, D35; 9: D33. Group D: 10: C*, C2, C16, C13, C17, I*, I3, I4, I16, I17, I25, I26; 11: I33, I31. Group E: 12: D20; 13: D29; 14: D32; 15: D11; 16: D10, D13; 17: D26; 18: D14; 19: D12; 20: D9. *WHO reference strains. Groups are as in Table XXII (insert).

Detailed results from tree construction:

- Distance methods
    - In phenograms from Jaccard distances four clusters of strains could be identified with all clustering methods (Fig. 27): UPGMA (CC = 0.99), neighbour joining and Fitch-Margoliash (SQ = 1.75; ASD = 5.9). The *L. donovani* complex was rooted between group D and the remainder.
- Parsimony methods:
    - Dollo produced the shortest cladograms (30 trees with 77 total steps). In the consensus tree (Fig. 27) the *L. donovani* complex was rooted between group AB and the remainder (BCD).
    - polymorphism generated 8 trees with 119 polymorphisms in each character and rooted the *L. donovani* complex between groups AB and CD, but with bootstrap support between 50 and 60%.
    - the consensus Wagner tree, from 20 trees with 176 total steps, had a topology similar to the clustering (distance) methods.
    - Camin-Sokal generated the longest cladograms (15 trees with 195 total steps), in which consensus tree group D was associated with group C, although with low bootstrap support.

Upon phylogenetic analysis of ITS RFLP it was possible to identify four major clusters of strains within the *L. donovani* complex (Fig. 27), which were congruent amongst distance methods. Parsimony methods were congruent, and congruent in relation to the clustering methods, for values higher than 75%, except for the root of the *L. donovani* complex in the Dollo tree. Dollo parsimony usually is the method of choice for restriction fragments and it has produced here the shortest trees, however, the Dollo cladogram was assymetric and suggested different evolutionary rates, contrary to all other methods.

Through the distance matrix in Table XXI definition of similarity triangles was poor, but many strains shared identical profiles. There was, however, a clear separation between group E and the remaining strains.

5.3. Mini-exon

5.3.1. PCR amplification

The mini-exon PCR products varied in size within the *L. donovani* complex, between 0.4 to 0.5kb, with double bands in the Indian *L. donovani* strains (Figure 28). However, the *L. donovani* complex could not be distinguished from other OW strains by size alone, despite some size polymorphism.

Figure 28 - Amplification products for miniexon. Reference strains: A - *L. aethiopica*; C - *L. chagasi*; I - *L. infantum*; D - *L. donovani*; M - *L. major*; T - *L. tropica*; a/b - larger and shorter fragments, respectively, from doublet amplification. For strain codes, see Tables II to V. The shorter fragment was also amplified with fragments Da and D4a.

5.3.2. Restriction analysis of mini-exon in the *Leishmania donovani* complex

Mini-exon PCR-RFLP profiles (Fig. 29) of *L. chagasi* were indistinguishable from most *L. infantum*, whilst *L. donovani* presented a high degree of diversity. The restriction fragments from the *L. donovani* complex were clearly distinct from the outgroup profiles. The full RFLP data are shown in Table XLI (Annex 2).

Total fragment size (Table XXII) varied from ~0.2 to ~1kb, whilst some strains had total sizes between the expected 0.4 to 0.5 kb.

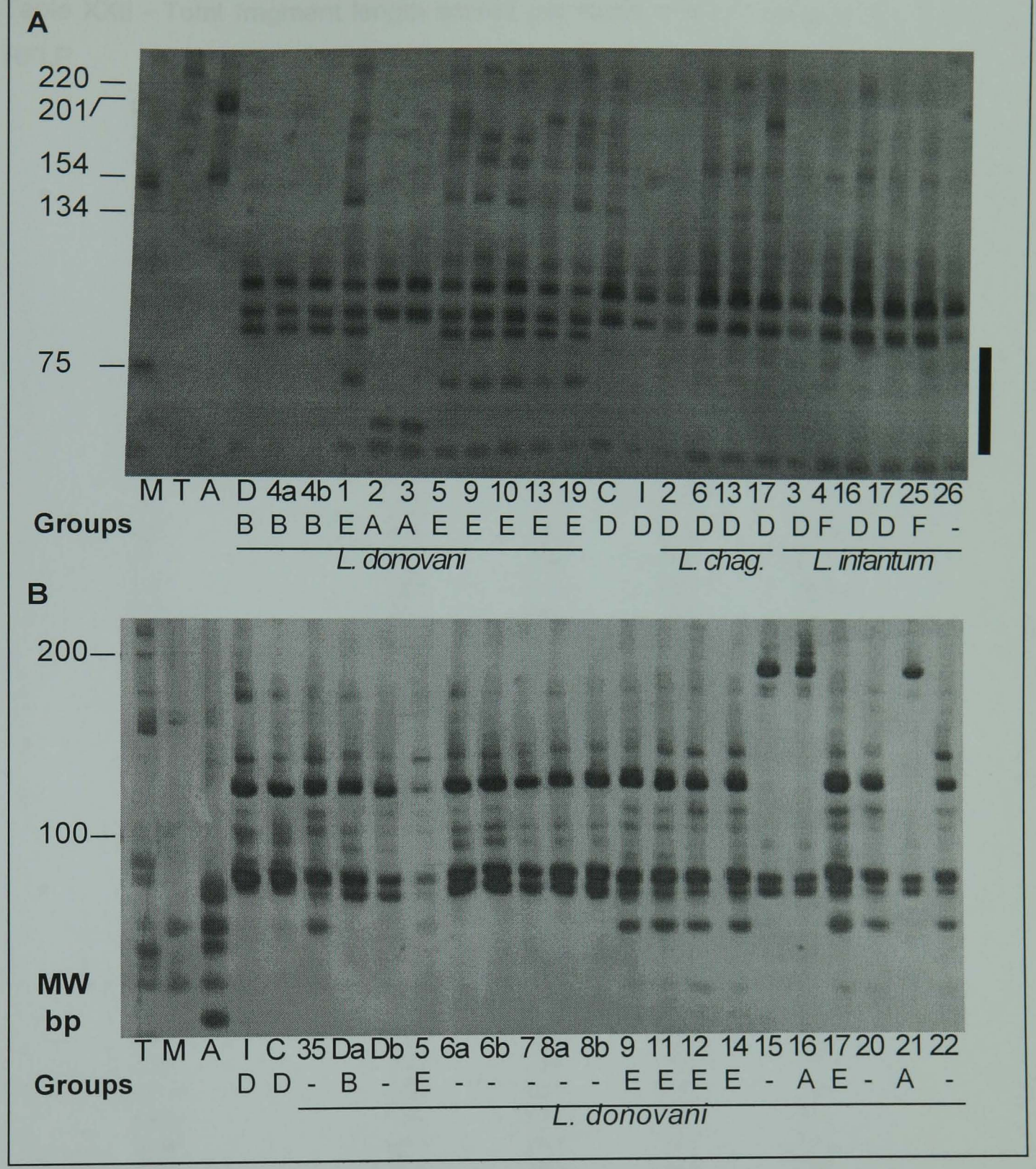


Figure 29 - Restriction products of miniexon within the *L. donovani* complex, with A) *Hae*III and B) *Msp*I. Absence of bands in the marked region is characteristic of *L. infantum* / *L. chagasi* strains. Genetic groups are shown for each profile.

Table XXII - Total fragment length scored per strain and per enzyme for mini exon RFLP.

Strain \ Enzyme	*BstUI*	*CfoI*	*HaeIII*	*MspI*
L. major	209	203	335	398
L. aethiopica	223	123	418	301
L. tropica	209	240	354	247
C	234	384	419	199
C2	234	244	253	199
C3	234	244	253	199
C6	234	244	253	199
C13	234	244	253	199
C17	234	244	253	199
C20	234	244	253	199
I	350	384	253	199
I3	234	244	253	199
I4	435	404	253	199
I16	234	244	253	346
I17	234	244	453	199
I25	551	404	328	346
D	234	244	340	277
D1*	886	479	990	346
D2	234	244	253	340
D3	234	244	253	340
D4	234	244	340	277
D4a	234	244	282	277
D5*	886	975	762	346
D6a	432	359	415	277
D6b	432	359	415	277
D7	432	359	415	277
D8a	432	359	415	277
D8b	432	359	415	277
D9*	842	760	415	329
D10*	730	760	908	346
D11*	842	760	415	329
D12*	842	760	415	329
D13*	551	760	762	346
D14*	842	760	415	329
D15	234	244	415	340
D16	234	384	315	340
D17*	954	851	315	329
D18*	749	975	928	346
D19*	886	893	928	346
D20	315	404	415	329
D21	234	244	315	340
D24*	234	544	402	329
D25	234	384	315	340
D26	234	384	315	340
D27*	234	625	315	340
D28*	234	625	415	329
D29	234	244	415	329
D30	234	384	315	340
D32	234	485	415	251
D33*	234	625	415	251
D34*	234	625	415	251
D35*	401	661	415	329

Fragment size was determined by comparison with molecular markers or extrapolation. Accumulation of small errors may cause some variation in total size differences, whilst heterogeneous targets may produce artificial larger sizes (*).

Table XXIII - Pairwise distance (Jaccard index) among *L. infantum* complex strains for mini-exon RFLP data. Values lower than 0.511 (65% of the maximum distance within the *L. donovani* complex) are shown in black on the upper half.

Group		M	A	T	C2	I17	C	I	I16	D	D4a	D29	D15	D16	D21	D2	D27	D6a	D28	D32	D33	D35	D20	D24	I25	I4	D5	D18	D19	D1	D10	D13	D11	D17
					D					B				A					C						F		E							
	M	-																																
	A	0.922	-																															
	T	0.827	0.882	-																														
D	C2	0.933	0.951	0.956	-																													
	I17	0.935	0.953	0.933	0.267	-																												
	C	0.938	0.956	0.959	0.365	0.433	-																											
	I	0.938	0.956	0.959	0.365	0.433	0.354	-																										
	I16	0.938	0.929	0.935	0.365	0.433	0.485	0.485	-																									
B	D	0.938	0.929	0.959	0.365	0.433	0.485	0.485	0.354	-																								
	D4A	0.959	0.926	0.957	0.447	0.500	0.542	0.542	0.433	0.258	-																							
	D29	0.920	0.935	0.941	0.485	0.527	0.562	0.562	0.471	0.343	0.420	-																						
	D15	0.917	0.933	0.961	0.542	0.577	0.607	0.607	0.527	0.420	0.485	0.408	-																					
A	D16	0.941	0.933	0.961	0.542	0.577	0.527	0.527	0.527	0.527	0.577	0.592	0.471	-																				
	D21	0.938	0.929	0.959	0.500	0.542	0.577	0.577	0.485	0.485	0.542	0.562	0.420	0.250	-																			
	D2	0.935	0.926	0.957	0.447	0.500	0.542	0.542	0.433	0.433	0.500	0.527	0.354	0.354	0.258	-																		
	D27	0.945	0.938	0.943	0.607	0.632	0.592	0.592	0.592	0.592	0.632	0.640	0.548	0.333	0.408	0.471	-																		
	D6A	0.903	0.917	0.945	0.562	0.592	0.617	0.548	0.548	0.459	0.513	0.447	0.500	0.640	0.617	0.592	0.677	-																
C	D28	0.928	0.943	0.926	0.592	0.617	0.577	0.577	0.577	0.500	0.548	0.387	0.535	0.603	0.640	0.617	0.522	0.552	-															
	D32	0.923	0.961	0.920	0.527	0.562	0.592	0.592	0.592	0.513	0.562	0.397	0.548	0.674	0.655	0.632	0.603	0.564	0.316	-														
	D33	0.926	0.962	0.923	0.562	0.592	0.548	0.548	0.617	0.548	0.592	0.447	0.577	0.640	0.674	0.655	0.564	0.590	0.224	0.229	-													
	D35	0.916	0.928	0.913	0.659	0.677	0.645	0.645	0.645	0.590	0.626	0.511	0.612	0.663	0.693	0.677	0.650	0.566	0.456	0.540	0.500	-												
	D20	0.928	0.943	0.947	0.592	0.617	0.640	0.640	0.577	0.500	0.548	0.387	0.535	0.659	0.640	0.617	0.693	0.552	0.511	0.522	0.552	0.361	-											
	D24	0.949	0.943	0.947	0.592	0.617	0.577	0.577	0.577	0.500	0.548	0.488	0.603	0.535	0.577	0.617	0.590	0.612	0.511	0.590	0.552	0.456	0.426	-										
F	I25	0.950	0.945	0.949	0.617	0.640	0.659	0.603	0.535	0.603	0.640	0.645	0.677	0.626	0.603	0.640	0.663	0.632	0.694	0.707	0.720	0.609	0.540	0.540	-									
	I4	0.943	0.959	0.962	0.485	0.527	0.562	0.562	0.562	0.562	0.607	0.617	0.655	0.655	0.632	0.607	0.692	0.659	0.677	0.640	0.659	0.577	0.488	0.564	0.436	-								
E	D5	0.920	0.930	0.918	0.771	0.750	0.759	0.729	0.729	0.729	0.750	0.718	0.739	0.767	0.786	0.778	0.728	0.637	0.651	0.696	0.674	0.577	0.651	0.686	0.628	0.685	-							
	D18	0.920	0.946	0.918	0.771	0.750	0.729	0.729	0.729	0.729	0.750	0.718	0.739	0.767	0.786	0.778	0.728	0.637	0.651	0.696	0.674	0.577	0.651	0.686	0.628	0.685	0.246	-						
	D19	0.920	0.930	0.918	0.771	0.750	0.729	0.729	0.729	0.729	0.750	0.718	0.739	0.767	0.786	0.778	0.728	0.637	0.651	0.696	0.674	0.618	0.686	0.717	0.664	0.718	0.246	0.246	-					
	D1	0.915	0.926	0.930	0.753	0.730	0.741	0.741	0.707	0.707	0.730	0.696	0.718	0.750	0.741	0.762	0.767	0.648	0.728	0.739	0.748	0.628	0.622	0.661	0.548	0.658	0.420	0.420	0.420	-				
	D10	0.915	0.943	0.930	0.753	0.762	0.741	0.741	0.707	0.707	0.730	0.696	0.718	0.778	0.771	0.762	0.739	0.606	0.661	0.672	0.685	0.586	0.622	0.696	0.596	0.658	0.478	0.420	0.478	0.535	-			
	D13	0.926	0.955	0.924	0.720	0.694	0.743	0.707	0.667	0.667	0.694	0.655	0.681	0.753	0.743	0.732	0.707	0.598	0.616	0.627	0.643	0.577	0.567	0.658	0.535	0.609	0.395	0.395	0.461	0.522	0.402	-		
	D11	0.928	0.921	0.926	0.732	0.743	0.753	0.719	0.719	0.681	0.707	0.627	0.695	0.762	0.753	0.743	0.718	0.567	0.587	0.598	0.616	0.491	0.535	0.632	0.606	0.627	0.485	0.535	0.535	0.586	0.433	0.475	-	
	D17	0.946	0.924	0.928	0.743	0.753	0.730	0.695	0.730	0.730	0.753	0.718	0.771	0.707	0.730	0.753	0.696	0.672	0.685	0.729	0.707	0.568	0.648	0.606	0.577	0.643	0.507	0.553	0.553	0.561	0.601	0.632	0.468	-

A is *L. aethiopica*, M is *L. major*, T is *L. tropica*. C2 = C3, C6, C13, C17, C20, I3; D = D4; D2 = D3; D6 = D7, D8; D10 = D9; D11 = D12, D14; D16 = D25, D26, D30.

Detailed results from tree construction:

- Distance methods
    - In the neighbour joining phenogram (Fig. 30), six groups were identified and the *L. donovani* complex was rooted between group A and the remainder.
    - by UPGMA (CC = 0.96) the phenogram was similar to neighbour joining but the branching order of B, D and A was different from either neighbour joining or Fitch-Margoliash.
    - by Fitch-Margoliash (SQ = 4.8; ASD = 6.2) produced the same groups as neighbour joining but the *L. donovani* complex was rooted between group B and the remainder.
- Parsimony methods:
    - Wagner (Fig. 30) generated the shortest cladograms (73 trees with 109 total steps). The same groups as by neighbour joining were produced, except that a clear group D was not resolved from a clade DF.
    - Dollo generated 26 trees with 115 total steps.
    - polymorphism parsimony generated 8 trees with 188 polymorphisms in each character.
    - Camin-Sokal generated 100 trees with 128 total steps.

In the phenograms built using distance methods (Fig. 30) six congruent groups were identified, although the branching order could be different (see above). Resolution was good for all groups except group D by Wagner, however affinities of several strains were not congruent between the different analyses. No major lineage within the *L. donovani* complex had bootstrap support higher than 70% and topology of consensus trees varied considerably among parsimony methods. Accordingly, in the distance matrix show in Table XXIII, there were no obvious groups formed.

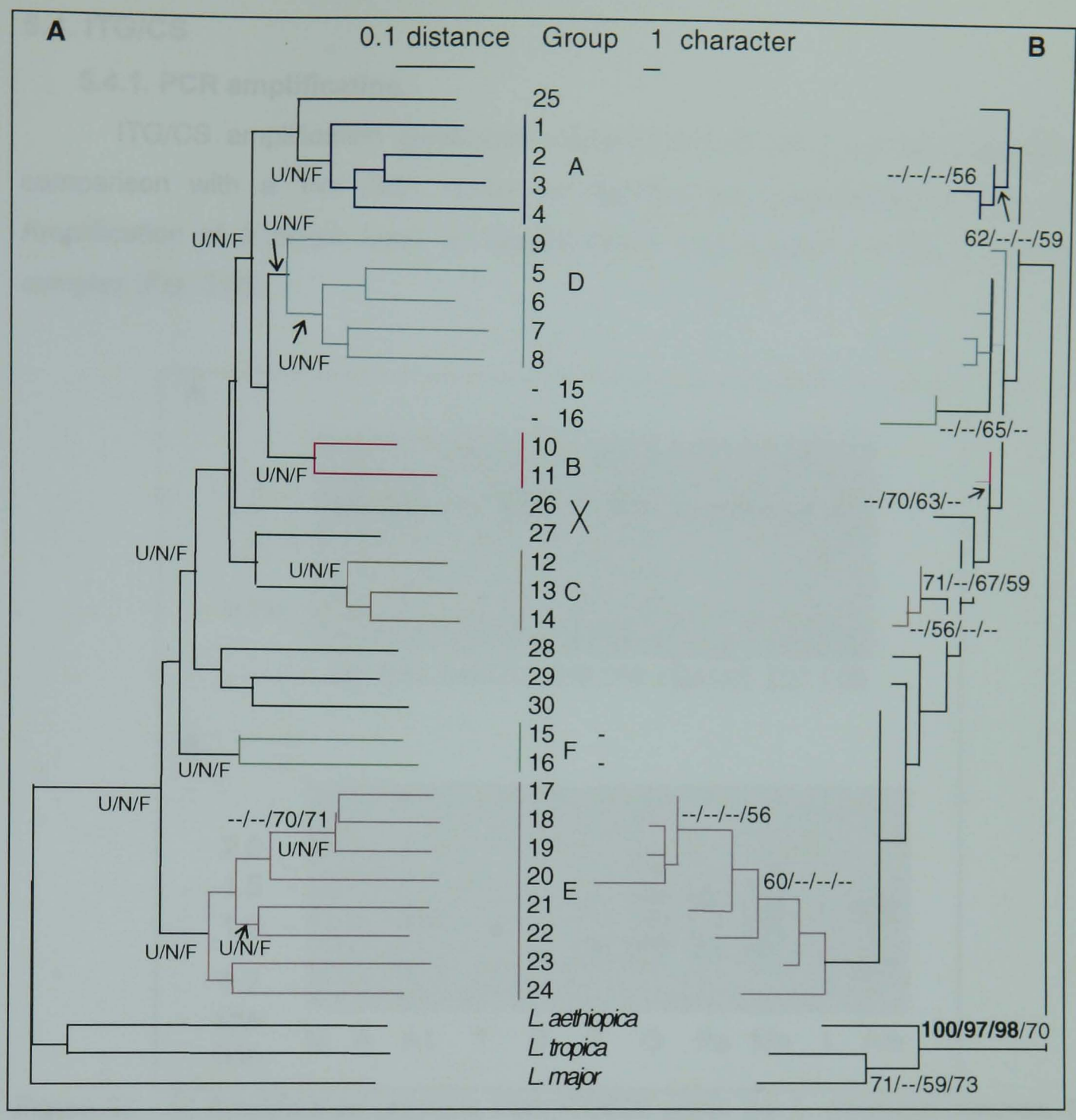


Figure 30 - Phylogenetic analysis of the *L. donovani* complex from mini-exon RFLP data. A) Neighbour-joining tree from Jaccard distances. Branches present in all three UPGMA, neighbour-joining and Fitch-Margoliash trees (U/N/F) are indicated. B) Wagner parsimony consensus tree. Bootstrap values higher than 60% (higher than 80% in bold) are shown for Wagner / Dollo / polymorphism / Camin-Sokal parsimony. Group A - 1: D2, D3; 2: D21; 3: D16, D25, D26, D30; 4: D27. Group D - 5: C2, C6, C13, C17, I3, C3, C20; 6: I17; 7: C*; 8: I*; 9: I16. Group B - 10: D*, D4; 11: D4a. Group C - 12: D28; 13: D33, D34; 14: D32. Group F - 15: I4; 16: I25. Group E - 17: D18; 18: D5; 19: D19; 20: D1; 21: D13; 22: D9, D10; 23: D11, D12, D14; 24: D17. Not grouped: 25: D15; 26: D29; 27: D6, D7, D8; 28: D24; 29: D20; 30: D35. * are reference strains. Groups are equivalent to those in Table XXXII (insert), except for F.

5.4. ITG/CS

5.4.1. PCR amplification

ITG/CS amplification generated a product of around 1.7kb as estimated by comparison with a 1kb DNA ladder on agarose gel electrophoresis (Fig. 31A) Amplification of a single band for gp63 ITG/CS was specific for the *L. donovani* complex (Fig. 31B).

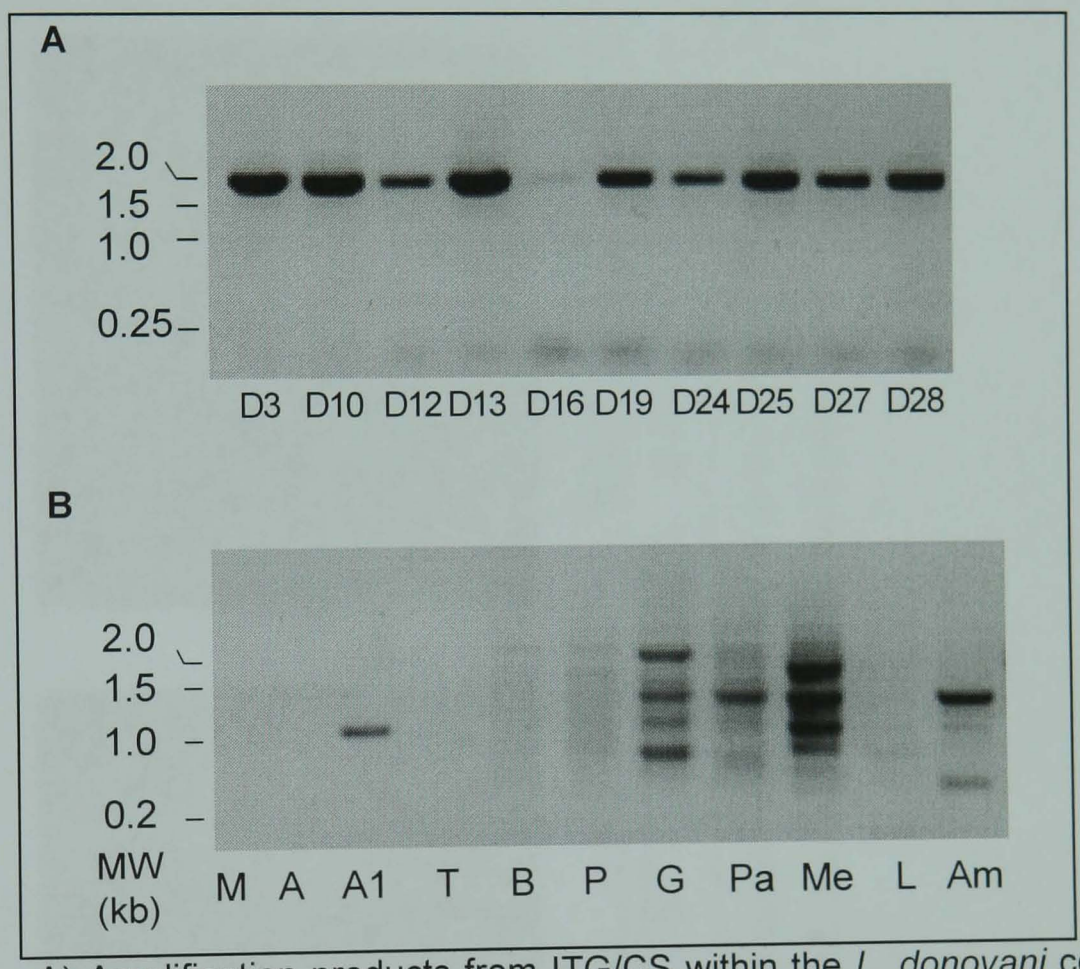


Figure 31 - A) Amplification products from ITG/CS within the *L. donovani* complex. B) Specificity of ITG/CS amplification. Reference strains: A - *L. aethiopica*, Am - *L. amazonensis*, B - *L. braziliensis*, G - *L. guyanensis*, L - *L. lainsoni*, M - *L. major*, Me - *L. mexicana*, P - *L. panamensis*, Pe - *L. peruviana*, T - *L. tropica*. MW are molecular weight markers.

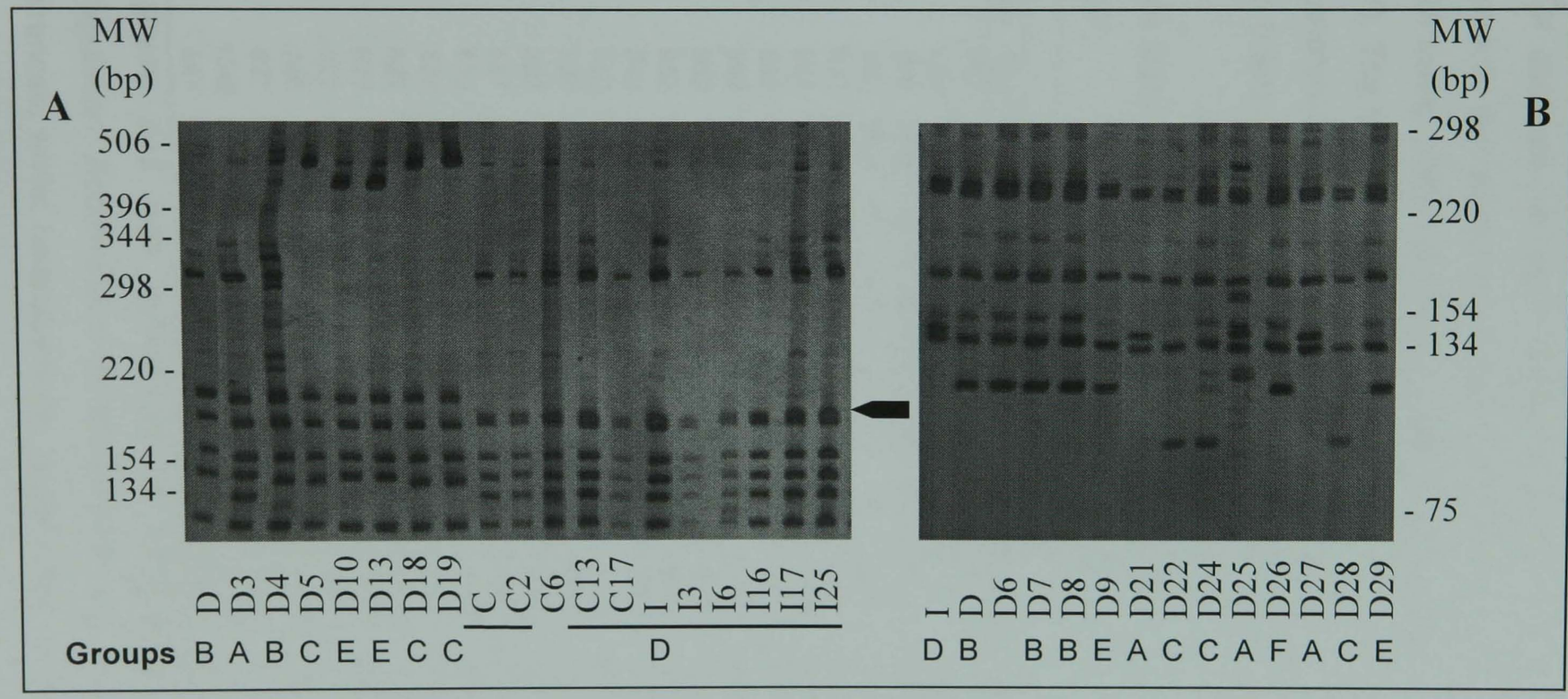


Figure 32 - Restriction profiles generated by enzymes A) *Cfo*I and B) *Hae*III on ITG/CS amplification products. *Cfo*I generated specific profiles for *L. infantum / L. chagasi* for the size indicated by an arrow. Genetic groups are shown for each profile, as in Table XXII (insert).

5.4.2. Restriction analysis of ITG/CS in the *Leishmania donovani* complex

Restriction of ITG/CS generated rich band profiles. Enzymes *Cfo*I (Fig. 32), *Alu*I and *Rsa*I, used upon gp63 ITG/CS produced profiles specific for *L. infantum* and *L. chagasi*, and most strains of these two species were indistinguishable. The full RFLP data are shown in Table XLIII (Annex 2). Only *Bst*UI and *Cfo*I produced fragments that added up to approximately the expected size (Table XXIV), but with *Bst*UI some strains had significantly larger total sizes (D16, D21, D25, D27, D30 and D31). The larger total sizes included strains with more fragments, thus indicating a polymorphic target. Enzyme *Rsa*I produced only 1 or 2 usable fragments (smaller than 0.9kb) and thus the scored total fragment size was very small.

Table XXIV - Total fragment length scored per strain and per enzyme for ITG/CS RFLP.

	Enzymes								Enzymes						
Strain	***Alu*I**	***Bst*UI**	***Cfo*I**	***Hae*III**	***Msp*I**	***Rsa*I**	***Taq*I**	**Strain**	***Alu*I**	***Bst*UI**	***Cfo*I**	***Hae*III**	***Msp*I**	***Rsa*I**	***Taq*I**
C	982	1605	1531	1025	1300	235	997	**D11**	1127	1486	1542	1000	1212	98	982
C2	982	1605	1531	1025	1300	235	997	**D12**	1132	1486	1544	1000	1212	98	427
C12	982	1605	1531	1025	1300	235	997	**D13**	947	1486	1544	1000	1212	98	982
C13	982	1605	1531	1025	1300	235	997	**D14**	1127	1486	1542	1000	1212	98	982
C17	982	1605	1531	1025	1300	235	997	**D15**	827	1600	1547	1027	1300	245	973
C20	982	1605	1531	1025	1300	235	997	**D16***	827	1996	1549	1027	1300	-	982
I	982	1605	1531	1025	1300	235	997	**D17**	827	1651	1527	980	1232	245	962
I3	982	1605	1531	1025	1300	235	997	**D18**	832	1651	1527	980	1232	245	952
I6	982	1605	1531	1025	1300	235	997	**D19**	832	1651	1527	980	1232	245	952
I16	982	1605	1531	1025	1300	235	997	**D20**	1127	1486	1542	1000	1212	98	982
I17	982	1605	1531	1025	1300	235	997	**D21***	740	1996	1549	1027	1300	-	992
I26	982	1605	1531	1025	1300	235	772	**D22**	827	1651	1524	980	1232	245	1177
I31	827	1600	1534	1027	1300	245	973	**D24**	832	1651	1527	980	1232	98	972
I33	982	1605	1531	1025	1300	235	997	**D25***	827	2076	1549	1027	1300	245	982
D	827	1561	1341	1000	1232	245	943	**D26**	992	1571	1406	1000	1232	245	982
D1	832	1576	1592	1127	1232	245	967	**D27***	745	1996	1549	1027	1517	245	982
D2	827	1535	1549	1027	642	245	982	**D28**	745	1651	1527	980	1232	98	967
D3	827	1600	1549	880	642	245	982	**D29**	1137	1486	1544	1000	1212	98	997
D4	827	1636	1342	1000	1232	245	742	**D30***	832	2076	1549	1027	1300	245	982
D5	827	1651	1528	980	1232	245	1153	**D31***	745	1996	1549	1027	1300	245	982
D7	740	1636	1341	1000	1232	98	943	**D32**	745	1486	1479	1000	1212	98	992
D8	827	1636	1341	1000	1232	98	943	**D33**	745	1651	1527	980	1232	245	967
D9	1045	1486	1479	1000	1212	98	410	**D34**	740	1651	1527	980	1232	245	969
D10	947	1486	1544	1000	1212	98	982	**D35**	827	1651	1527	980	1232	98	952

Fragment size was determined by comparison with molecular markers or extrapolation. Accumulation of small errors may cause some variation in total size differences, whilst heterogeneous targets may produce total sizes larger than the PCR products (*).

Table XXV - Pairwise distance (Jaccard index) among *L. donovani* complex strains for ITG/CS RFLP data. Values lower than 0.502 (65% of the maximum distance within the *L. donovani* complex) are shown in black on the upper half.

Group		B				F	C										A										E						D	
		D7	D8	D	D4	D26	D17	D34	D35	D5	D28	D33	D22	D24	D1	D18	D27	D2	D3	D16	D25	D30	D31	D15	D21	I31	D9	D32	D12	D10	D29	D11	C	I26
B	D7	-																																
	D8	0.1-67	-																															
	D	0.285	0.232	-																														
	D4	0.363	0.324	0.324	-																													
F	D26	0.512	0.488	0.488	0.488	-																												
C	D17	0.494	0.469	0.469	0.469	0.544	-																											
	D34	0.474	0.494	0.494	0.494	0.528	0.281	-																										
	D35	0.474	0.447	0.494	0.494	0.528	0.281	0.232	-																									
	D5	0.512	0.488	0.488	0.442	0.558	0.354	0.320	0.320	-																								
	D28	0.500	0.518	0.556	0.556	0.550	0.392	0.363	0.363	0.469	-																							
	D33	0.518	0.535	0.535	0.535	0.528	0.358	0.324	0.392	0.442	0.167	-																						
	D22	0.550	0.528	0.528	0.488	0.558	0.354	0.320	0.320	0.316	0.418	0.387	-																					
	D24	0.518	0.494	0.535	0.535	0.528	0.358	0.392	0.324	0.442	0.285	0.324	0.442	-																				
	D1	0.584	0.564	0.528	0.564	0.522	0.413	0.442	0.442	0.482	0.358	0.320	0.436	0.387	-																			
	D18	0.535	0.512	0.512	0.512	0.463	0.320	0.281	0.281	0.354	0.324	0.281	0.354	0.281	0.354	-																		
A	D27	0.626	0.636	0.636	0.608	0.537	0.670	0.636	0.661	0.655	0.652	0.636	0.655	0.661	0.629	0.619	-																	
	D2	0.632	0.615	0.615	0.584	0.544	0.626	0.615	0.615	0.608	0.684	0.668	0.608	0.668	0.636	0.626	0.383	-																
	D3	0.603	0.584	0.584	0.550	0.544	0.626	0.615	0.615	0.608	0.684	0.668	0.608	0.668	0.661	0.626	0.383	0.229	-															
	D16	0.615	0.596	0.626	0.596	0.590	0.661	0.652	0.626	0.645	0.692	0.700	0.645	0.677	0.670	0.661	0.408	0.413	0.413	-														
	D25	0.619	0.601	0.601	0.571	0.565	0.612	0.601	0.601	0.595	0.670	0.655	0.595	0.655	0.623	0.612	0.374	0.378	0.378	0.271	-												
	D30	0.645	0.629	0.629	0.601	0.532	0.639	0.629	0.629	0.623	0.645	0.629	0.623	0.629	0.595	0.583	0.309	0.431	0.431	0.345	0.218	-											
	D31	0.626	0.636	0.636	0.608	0.537	0.670	0.636	0.661	0.655	0.652	0.636	0.655	0.661	0.629	0.619	0.224	0.436	0.436	0.349	0.309	0.221	-										
	D15	0.596	0.577	0.577	0.544	0.601	0.645	0.661	0.661	0.601	0.700	0.685	0.655	0.685	0.655	0.670	0.477	0.436	0.436	0.408	0.374	0.426	0.431	-									
	D21	0.603	0.615	0.643	0.615	0.636	0.652	0.668	0.668	0.685	0.684	0.692	0.685	0.692	0.685	0.700	0.436	0.488	0.488	0.277	0.378	0.431	0.383	0.436	-								
	I31	0.643	0.626	0.626	0.626	0.645	0.636	0.652	0.652	0.619	0.692	0.677	0.645	0.677	0.645	0.661	0.500	0.463	0.463	0.436	0.403	0.452	0.457	0.349	0.463	-							
E	D9	0.584	0.598	0.598	0.598	0.650	0.640	0.528	0.629	0.650	0.707	0.714	0.727	0.714	0.676	0.667	0.648	0.692	0.692	0.639	0.616	0.589	0.623	0.595	0.655	0.639	-						
	D32	0.617	0.629	0.629	0.629	0.591	0.640	0.657	0.657	0.676	0.617	0.629	0.676	0.629	0.650	0.640	0.632	0.667	0.667	0.700	0.700	0.677	0.659	0.728	0.667	0.722	0.465	-					
	D12	0.610	0.591	0.591	0.591	0.584	0.632	0.650	0.622	0.643	0.667	0.676	0.643	0.650	0.643	0.659	0.677	0.632	0.632	0.668	0.670	0.693	0.700	0.700	0.684	0.692	0.333	0.459	-				
	D10	0.676	0.659	0.659	0.659	0.528	0.692	0.684	0.659	0.677	0.676	0.684	0.677	0.659	0.652	0.643	0.636	0.643	0.643	0.677	0.678	0.655	0.661	0.728	0.714	0.721	0.480	0.429	0.363	-			
	D29	0.659	0.643	0.643	0.643	0.544	0.677	0.692	0.668	0.685	0.632	0.643	0.661	0.643	0.608	0.652	0.645	0.652	0.652	0.685	0.686	0.663	0.670	0.714	0.700	0.707	0.453	0.397	0.324	0.281	-		
	D11	0.684	0.668	0.668	0.643	0.544	0.700	0.692	0.668	0.661	0.684	0.692	0.685	0.668	0.661	0.652	0.670	0.677	0.677	0.707	0.707	0.686	0.693	0.714	0.741	0.747	0.541	0.500	0.447	0.281	0.387	-	
D	C	0.747	0.734	0.752	0.714	0.663	0.740	0.752	0.734	0.727	0.721	0.728	0.740	0.707	0.707	0.721	0.639	0.655	0.655	0.629	0.606	0.577	0.612	0.583	0.645	0.629	0.772	0.735	0.728	0.693	0.678	0.655	-	
	I26	0.741	0.728	0.747	0.707	0.655	0.734	0.747	0.728	0.721	0.647	0.657	0.650	0.657	0.721	0.714	0.707	0.645	0.645	0.722	0.721	0.741	0.728	0.748	0.715	0.742	0.767	0.735	0.721	0.685	0.693	0.645	0.158	-

D10 = D13; D18 = D19; C =C2, C12, C13, C17, I, I3, I6, I16, I17, I33, C20; D11 = D14, D20; D34 = D35.

Detailed results from tree construction:

- Distance methods
    - The neighbour joining, UPGMA (CC = 0.98) and Fitch-Margoliash (SQ = 3.7; ASD = 5.9) phenograms had similar main topologies and the same groups of strains were identified (Fig. 33).
- Parsimony methods:
    - Dollo generated 39 trees with 136 total steps, in which consensus group E was closer to group D than group A, whilst the remainder were similar;
    - Wagner (Fig. 30) generated the shortest parsimony trees (96 trees with 117 total steps), for which the consensus had the same groups and main topology as the neighbour joining tree;
    - polymorphism generated 4 trees with 233 polymorphisms in each character, in which consensus strain D26 (Wangjie 1) clustered with group B, although with low bootstrap support, and group C was further sub-divided in two (strains 1-5 and 6-10);
    - Camin-Sokal generated 20 alternative trees with 181 total steps, in which consensus, strains I31 (Buck) and D15 (Addis 142) clustered with group D (*L. infantum*) instead of group A, although at a large distance (a total of 17 steps).

Phylogenetic analyses of RFLP of gp63 ITG/CS produced very consistent results across the different methods used. The only exceptions were the exchange of positions of groups D and A in the Dollo cladogram, clustering of strains I31 and D15 with *L. infantum* in the Camin-Sokal cladogram and the relative positions of strain Wangjie 1 (D26) which could be between CB and ADE or near group B using different methods. Strain groups B, D and E had the highest boostrap support, but neither group A nor the relative positions of the groups were well supported.

In the gp63 ITG/CS PCR-RFLP distance matrix in Table XXV, five groups of closely related strains could be identified visually, which had been defined in the phylogenetic analysis by all clustering methods (Fig. 33).

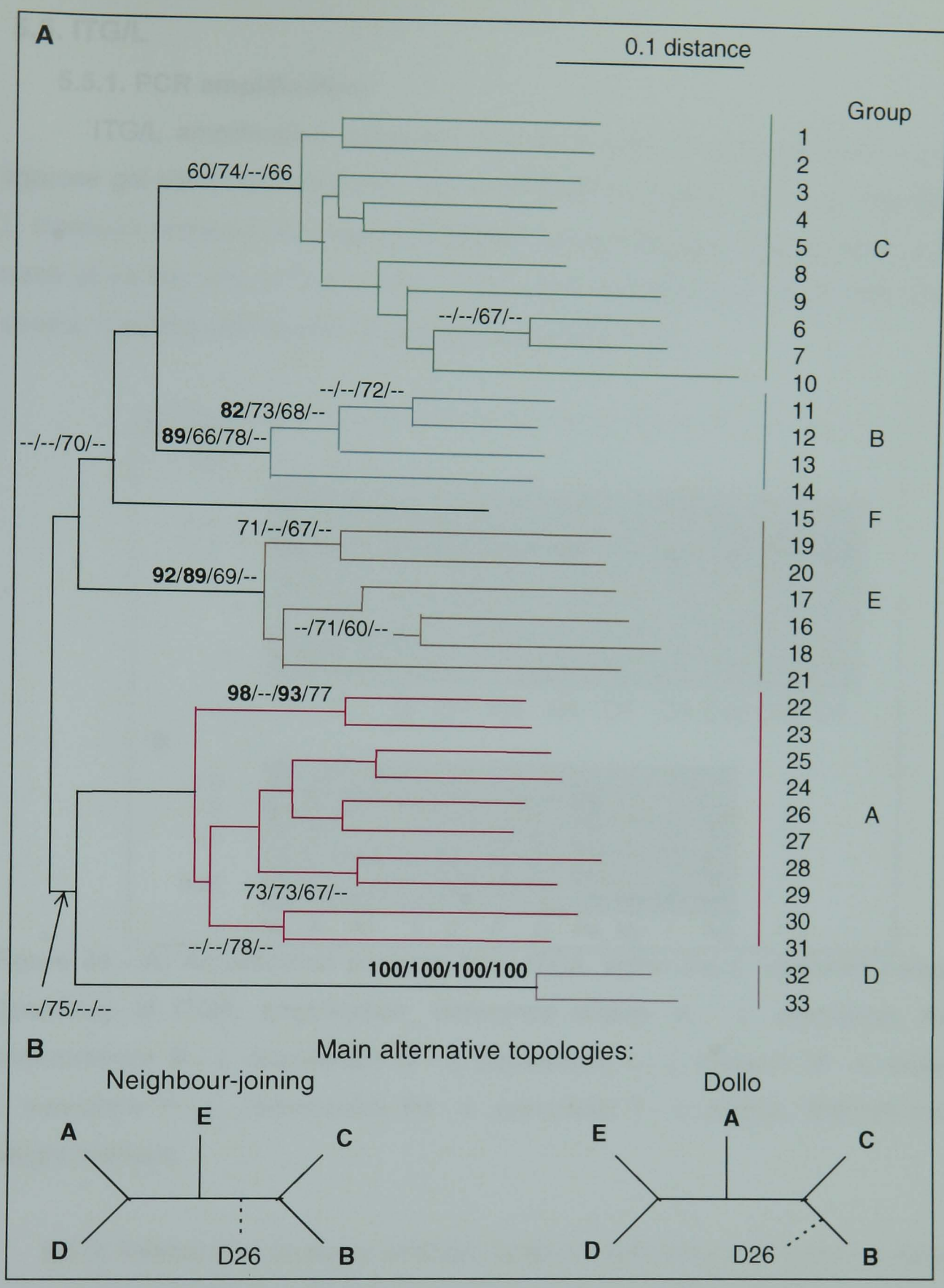


Figure 33 - Phylogenetic analysis of the *L. donovani* complex from ITG/CS RFLP data. A) Neighbour-joining tree, with topology compared to a Dollo cladogram. B) Main alternative topologies. Bootstrap support higher than 60% (higher than 80% in bold) is shown for parsimony methods (Dollo / Wagner / polymorphism / Camin-Sokal) at branches. Group C - 1: D5; 2: D22; 3: D17; 4: D35; 5:D34; 6: D28; 7: D33; 8: D18, D19; 9: D24; 10: D1. Group B - 11: D7; 12: D8; 13: D*; 14: D4. Group F - 15: D26. Group E - 16: D10, D13; 17: D29; 18: D11, D14, D20; 19: D12; 20: D9; 21: D32. Group A - 22: D2; 23: D3; 24: D25; 25: D30; 26: D31; 27: D27; 28: D16; 29: D21; 30: I31; 31: D15. Group D - 32: C*, C2, C12, C13, C17, I*, I3, I6, I16, I17, I33, C20; 33: I26. * WHO reference strains. Groups are the same as those in Table XXXII (insert).

5.5. ITG/L

5.5.1. PCR amplification

ITG/L amplification generated a product with an estimated size of 1.6kb on agarose gel electrophoresis (Fig. 34). Amplification of gp63 ITG/L was specific for the *L. donovani* complex, although at high DNA concentrations it was possible to amplify a band of similar size to that of the *L. donovani* complex from some OW *Leishmania* strains, together with many non-specific bands (Fig. 34).

Figure 34 - A) Amplification products from ITG/L within the *L. donovani* complex. B) Specificity of ITG/L amplification. Reference strains: A - *L. aethiopica*, Am - *L. amazonensis*, B - *L. braziliensis*, G - *L. guyanensis*, L - *L. lainsoni*, M - *L. major*, Me - *L. mexicana*, P - *L. panamensis*, Pe - *L. peruviana*, T - *L. tropica*. MW are molecular weight markers.

5.5.2. Restriction analysis of ITG/L in the *Leishmania donovani* complex

The restriction profiles of *L. infantum* and *L. chagasi* strains were identical, except for strains Lombardi and Strain A (Fig. 35) whilst diversity within *L. donovani* was evident. Most strains had similar profiles and few polymorphisms were detected within the *L. donovani* complex compared with ITG/CS. The full RFLP data are shown in Table XLV Annex 2.

The total usable fragment size for ITG/L (Table XXVI) was usually smaller than the expected and ranged between 1.58kb (some strains with *Msp*I) and 185bp (*Taq*I). Whilst with some enzymes larger bands were not scored, some fragments must have had the same size and thus have been indistinguishable with other enzymes. ITG/L RFLP profiles were often very complex, with small bands produced, and thus the opportunity for band coincidence was greater.

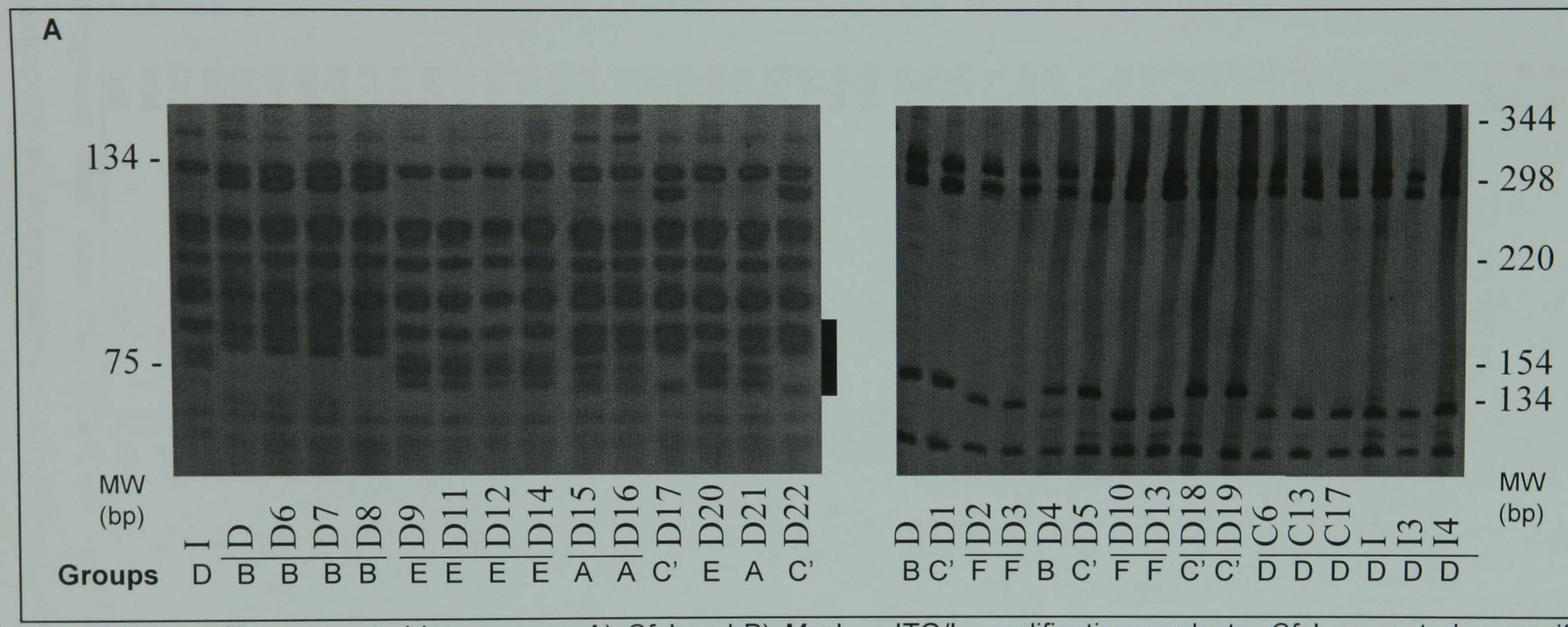


Figure 35 - Restriction profiles generated by enzymes A) *Cfo*I and B) *Msp*I on ITG/L amplification products. *Cfo*I generated a specific profile for *L. infantum*. The restriction patterns are also very different from those obtained with ITG/CS. A vertical bar highlights region of *L. infantum* specific band patterns. The groups corresponding to each profile are shown and are equivalent for most strains to those in Table XXXII (insert), except for group F.

Table XXVI - Total fragment length per strain and per enzyme for ITG/L RFLP.

Strain \ Enzyme	*AluI*	*BstUI*	*CfoI*	*HaeIII*	*MspI*	*RsaI*	*TaqI*
C	681	845	598	959	887	1407	185
C2	681	845	598	959	887	1407	185
C6	681	845	598	959	887	1407	185
C13	681	845	598	959	887	1407	185
C17	681	845	598	959	887	1407	185
C20	681	965	520	959	1032	1407	185
I	681	845	598	959	887	1407	185
I3	681	845	598	959	887	1407	185
I4	681	845	598	959	887	1407	185
I16	681	845	598	959	887	1407	185
I17	681	845	598	959	887	1407	185
I25	681	845	598	959	887	1407	185
I26	681	845	598	959	887	1407	185
I31	681	845	745	959	1032	1407	185
I33	681	975	673	959	1032	1407	185
D	681	1074	727	959	899	1407	185
D1	681	1148	724	959	899	1393	185
D2	681	948	675	959	891	1407	185
D3	681	948	675	959	891	1407	185
D4	681	1074	727	959	899	1407	185
D5	681	1278	724	959	899	1393	185
D6	681	1074	727	959	1146	1407	185
D7	681	1074	727	959	1146	1407	185
D8	681	1074	727	959	1146	1407	185
D9	681	845	745	959	1536	1407	185
D10	681	845	725	959	887	1407	185
D11	681	845	745	959	1536	1407	185
D12	681	845	745	959	1021	1407	185
D13	681	845	725	959	887	1407	185
D14	681	845	745	959	1021	1407	185
D15	681	1198	745	1166	1559	1407	185
D16	681	1198	745	1166	1559	1407	185
D17	681	1278	794	959	1159	1393	185
D18	681	1278	644	959	899	1393	185
D19	681	1148	644	959	899	1393	185
D20	681	845	745	959	1026	1407	185
D21	681	948	745	1166	1575	1407	185
D22	681	1148	794	959	894	1393	185
D24	681	1278	871	959	894	1393	185
D25	681	1198	747	1166	1575	1407	185
D26	681	1095	745	959	1186	1407	185
D27	681	1198	670	1166	1575	1407	185
D28	681	1278	871	959	1158	1362	185
D29	681	971	745	959	1036	1407	185
D30	681	948	747	1166	1045	1362	185
D31	681	948	747	1166	1580	1362	185
D32	681	845	745	959	1302	1362	185
D33	681	1198	871	959	1179	1362	185
D34	681	1278	794	959	894	1362	185
D35	681	1045	724	959	1041	1393	185

Fragment size was determined by comparison with molecular markers or extrapolation. Accumulation of small errors may cause some variation in total size differences, whilst fragments of the same size may cause artificial smaller sizes.

Table XXVII - Pairwise distance (Jaccard index) among *L. donovani* complex strains for ITG/L RFLP data. Values lower than 0.413 (65% of the maximum distance within the *L. donovani* complex) are shown in black on the upper half.

Group	Strain	C	C20	D10	I33	D9	D12	D26	D29	D32	D2	D	D6	D21	D15	D25	D27	D1	D5	D22	D18	D19	D17	D34	D24	D28	D30	D31	D33
D	C	-																											
	C20	0.297	-																										
	D10	0.174	0.338	-																									
E	I33	0.293	0.333	0.333	-																								
	D9	0.368	0.397	0.397	0.324	-																							
	D12	0.293	0.333	0.333	0.236	0.232	-																						
	D26	0.397	0.363	0.424	0.281	0.354	0.281	-																					
	D29	0.333	0.429	0.368	0.363	0.358	0.285	0.320	-																				
	D32	0.459	0.480	0.480	0.424	0.418	0.363	0.442	0.447	-																			
F	D2	0.408	0.435	0.435	0.429	0.474	0.429	0.494	0.453	0.541	-																		
B	D	0.435	0.459	0.403	0.500	0.535	0.500	0.512	0.424	0.591	0.368	-																	
	D6	0.459	0.480	0.429	0.518	0.550	0.518	0.528	0.447	0.571	0.397	0.167	-																
A	D21	0.518	0.535	0.535	0.528	0.558	0.528	0.537	0.506	0.577	0.358	0.442	0.413	-															
	D15	0.550	0.528	0.564	0.522	0.552	0.558	0.494	0.537	0.629	0.413	0.482	0.500	0.309	-														
	D25	0.550	0.528	0.564	0.522	0.552	0.522	0.452	0.500	0.571	0.413	0.482	0.457	0.309	0.305	-													
	D27	0.535	0.512	0.550	0.544	0.601	0.577	0.516	0.558	0.619	0.442	0.463	0.436	0.271	0.267	0.267	-												
C	D1	0.506	0.480	0.524	0.556	0.584	0.556	0.528	0.535	0.571	0.453	0.424	0.447	0.506	0.500	0.500	0.482	-											
	D5	0.524	0.500	0.541	0.535	0.596	0.571	0.506	0.550	0.584	0.474	0.447	0.469	0.522	0.477	0.477	0.457	0.164	-										
	D22	0.524	0.500	0.541	0.571	0.596	0.571	0.544	0.550	0.584	0.474	0.447	0.469	0.482	0.477	0.516	0.457	0.164	0.229	-									
	D18	0.506	0.480	0.524	0.518	0.584	0.556	0.488	0.535	0.571	0.500	0.474	0.494	0.544	0.500	0.500	0.482	0.232	0.164	0.281	-								
	D19	0.487	0.459	0.506	0.541	0.571	0.541	0.512	0.518	0.556	0.480	0.453	0.474	0.528	0.522	0.522	0.506	0.167	0.232	0.232	0.167	-							
	D17	0.556	0.535	0.571	0.564	0.619	0.596	0.537	0.577	0.608	0.512	0.488	0.506	0.477	0.426	0.471	0.403	0.277	0.226	0.226	0.277	0.320	-						
	D34	0.577	0.556	0.591	0.584	0.636	0.615	0.558	0.596	0.564	0.535	0.512	0.528	0.537	0.494	0.532	0.477	0.320	0.277	0.277	0.320	0.358	0.274	-					
	D24	0.556	0.535	0.571	0.528	0.558	0.528	0.457	0.506	0.544	0.469	0.488	0.506	0.516	0.471	0.426	0.494	0.277	0.226	0.316	0.277	0.320	0.312	0.349	-				
	D28	0.603	0.584	0.615	0.577	0.601	0.577	0.516	0.558	0.522	0.528	0.544	0.558	0.532	0.489	0.447	0.511	0.383	0.349	0.408	0.383	0.413	0.345	0.312	0.271	-			
	D30	0.563	0.577	0.577	0.571	0.564	0.535	0.544	0.512	0.469	0.424	0.494	0.512	0.436	0.477	0.431	0.500	0.469	0.488	0.488	0.512	0.494	0.522	0.463	0.436	0.408	-		
	D31	0.591	0.603	0.603	0.596	0.590	0.564	0.571	0.544	0.463	0.469	0.528	0.506	0.378	0.471	0.374	0.452	0.506	0.522	0.522	0.544	0.528	0.516	0.500	0.477	0.403	0.226	-	
	D33	0.591	0.571	0.603	0.596	0.590	0.564	0.537	0.506	0.506	0.512	0.488	0.506	0.552	0.546	0.511	0.565	0.413	0.436	0.436	0.463	0.442	0.477	0.408	0.378	0.345	0.383	0.431	-

C = C2, C6, C13, C17, I, I3, I4, I16, I17, I25, I26; D = D4; D2 = D3; D6 = D7, D8; D9 = D11; D10 = D13; D12 = D14, D20, I31; D15 = D16; D34 = D35.

Detailed results from tree construction:

- seven clusters of strains were identified across the different phylogenetic methods used.
- Distance methods
    - the neighbour joining tree (Fig. 36A) assuming a molecular clock, is almost star like in shape with four main clusters: C', C"A, BF and ED.
    - in the UPGMA tree (CC = 0.95) was similar to the neighbour joining tree except that group C' was closer to cluster ED.
    - a Fitch-Margoliash tree (SQ = 5.0; ASD = 8.1) was very similar to the neighbour joining tree, but groups B and F did not cluster together.
- Parsimony methods:
    - the consensus Dollo tree (34 trees, 80 total steps) was similar to the neighbour joining tree except that strains D2 and D3 (#12, group F) were included within group A. Although group D was identified, some strains (#6 and 7) from group E were closer to group D.
    - the consensus Wagner tree (100 trees, 69 total steps) (Fig. 36B) was star like in shape and had group C" clustered with group C' and group A with F.
    - the consensus polymorphism tree (9 trees, 133 total polymorphisms) had a similar topology to the neighbour joining tree.
    - the consensus Camin-Sokal tree (100 trees, 109 total steps) included strain D32 (# 9) in group C and disrupted group D to be scattered in the root of group E. The method placed the root between group C and the remainder.

In the gp63 ITG/L PCR-RFLP trees (Fig. 36) seven clusters of strains were identified. Groups E and D were always associated but the relationships between other groups were variable (Fig. 63C). The weaknesses of the phylogenies were reflected on the bootstrap values, which were only consistently higher than 70% for groups C" and B.

In the Jaccard distance matrix (Table XXVII) some groups were easily identified, although other clusters were not obvious, and distances were generaly small, just over 0.6.

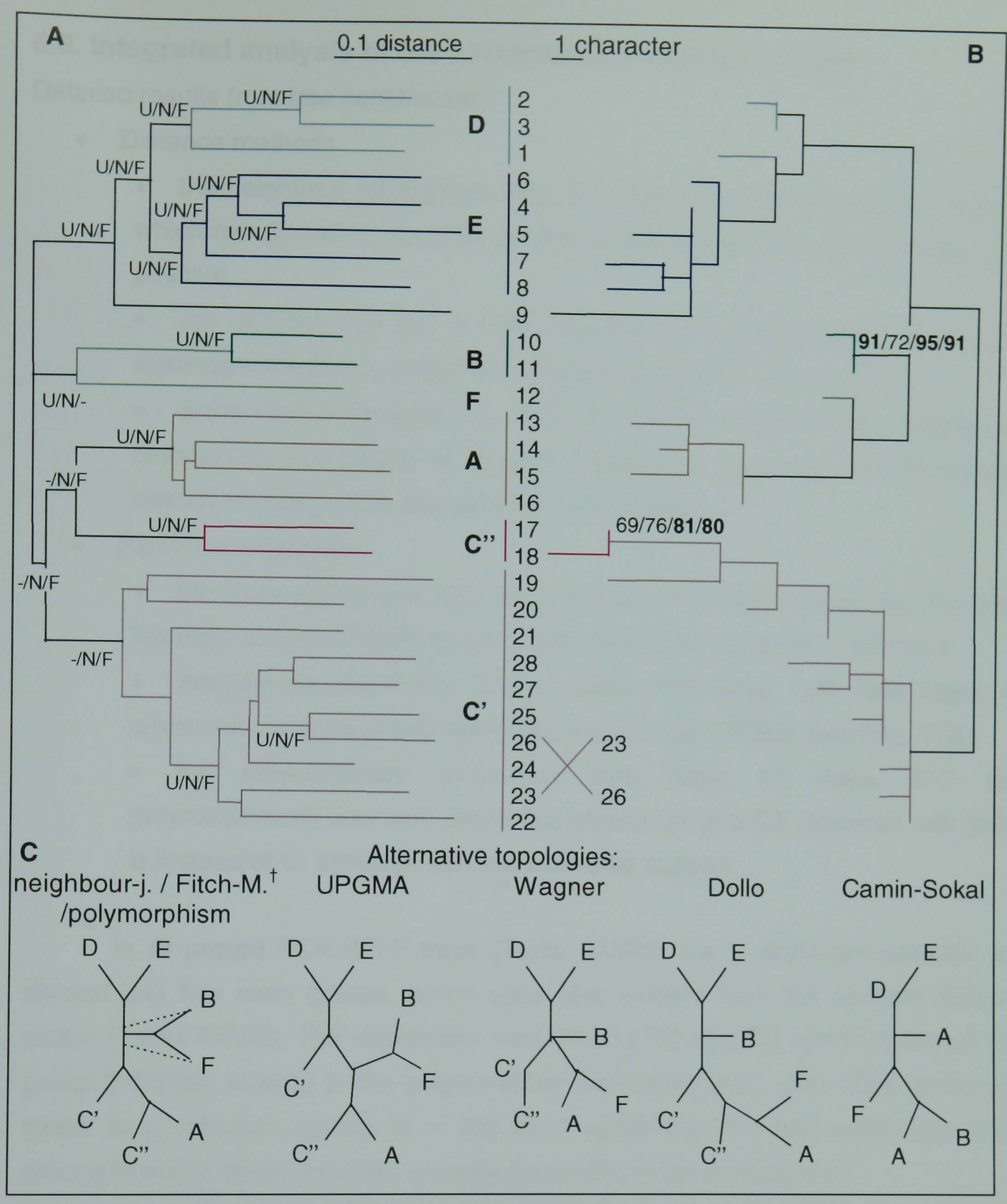


Figure 36 - Phylogenetic analysis of the *L. donovani* complex from ITG/L RFLP data with alternative topologies (C). A) Neighbour-joining tree. Branches present in UPGMA / neighbour-joining / Fitch-Margoliash (U/N/F) trees are indicated. B) Consensus Wagner parsimony tree. Bootstrap values higher than 69% (higher than 80% in bold) are shown for Wagner / Dollo / polymorphism / Camin-Sokal parsimony. Group D: 1 - C20, 2 - C*, C2, C6, C13, C17, I*, I3, I4, I16, I17, I25, I26; 3 - D10, D13. Group E: 4 - I31, D12, D14, D20; 5 - D9, D11; 6 - I33; 7 - D26; 8 - D29. Ungrouped: 9 - D32. Group B: 10 - D*, D4; 11 - D6, D7, D8. Group F: 12 - D2, D3. Group A: 13 - D25; 14 - D27; 15 - D15, D16; 16 - D21. Group C'': 17 - D30; 18 - D31. Group C': 19 - D33; 20 - D28; 21 - D24; 22 - D1; 23 - D5; 24 - D22; 25 - D18; 26 - D19; 27 - D17; 28 - D34, D35. * WHO reference strains. † dotted line indicates alternative branching. Groups are equivalent to those in Table XXXII (insert) for most strains, except group F.

5.6. Integrated analysis of the *Leishmania donovani* complex

Detailed results from tree construction:

- Distance methods
  - the neighbour joining tree (Fig. 37A) had five main groups of strains, of which two (E and C) could be divided in sub-groups. All branch lengths were positive.
  - the UPGMA tree (CC = 0.97) had the same groups and a very similar topology throughout, except that group A was closer to group B.
  - in the Fitch-Margoliash tree (SQ = 2.7; ASD = 4.4) the main clusters were maintained, but groups A, B and C formed a tricotomy, due to negative branch lengths, which also ocurred within groups.
- Parsimony methods:
  - the Dollo consensus tree (from 6 trees, 490 total steps) had the same topology as the UPGMA tree, but with strain D26 closer to *L. infantum*.
  - Wagner produced the shortest trees (100 trees, 395 total steps) for which the consensus had similar topology to the UPGMA tree (Fig. 37B).
  - the polymorphism consensus tree (from 16 trees, 812 total polymorphisms) was very similar but strains D2 and D3 clustered with group B instead of C, although with low bootstrap support.

In all pooled PCR-RFLP trees (Table XXVIII), the *L. donovani* complex was divided into five main groups, which were also evident from the Jaccard distance matrix (Table XXVIII). The exceptions were strains D2 and D3 which clustered with group B (Indian strains) in the polymorphism parsimony tree, strain D26 which was closer to *L. infantum* (group D) in the Dollo parsimony tree and subgroups within groups C and E which were not strongly supported by parsimony trees.

Topologies were very similar (Fig. 36C) with the exceptions of the neighbour joining tree in which group B clustered with group C instead of group A and Fitch-Margoliash which did not resolve the node between groups A, B and C.

Bootstrap support for the main groups of strains and nodes was higher than 90% by Wagner parsimony, except node A-B, followed by Dollo parsimony for which group A was not robust (54%). Strains from group A frequently associated with Indian *L. donovani* (group B) most notably in ITS RFLP trees. Low diversity groups such as *L. infantum* (group D) and Indian *L. donovani* (group B) were the most robust, as did cluster ABC. Afinities of strain D26 to group E were only distant as, in the consensus Dollo tree, it was closer to *L. infantum* with a bootstrap support of 73%.

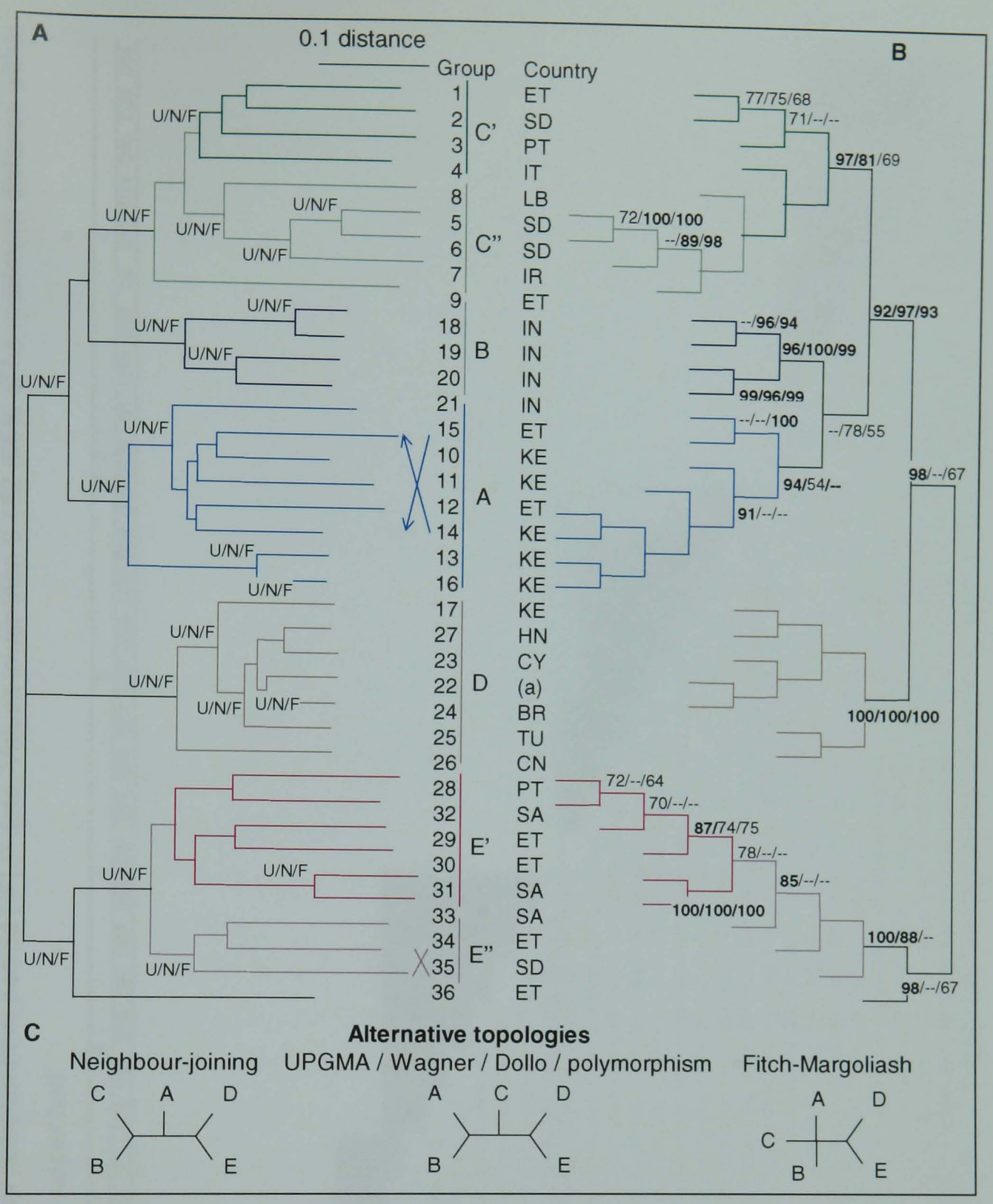


Figure 37 - Phylogenetic analysis of the *L. donovani* complex using pooled RFLP data, with alternative topologies (C). A) Neighbour-joining tree. Branches present in all UPGMA, neighbour-joining and Fitch-Margoliash trees (U/N/F) are indicated. B) Consensus Wagner parsimony tree. Bootstrap values higher than 50% (in bold higher than 80%) from 100 replicate data sets are shown for Wagner / Dollo / polymorphism parsimony. Groups C' - 1: D28; 2: D33; 3: D34; 4: D24. C" - 5: D18; 6: D19; 7: D5; 8: D17; 9: D1. A - 10: D16; 11: D25; 12: D27; 13: D30; 14: D21; 15: D15; 16: D2; 17: D3. B - 18: D7; 19: D8; 20: D*; 21: D4. D - 22 (a): C2 (PA), C13 (BR), C17 (HN), I3 (FR); 23: I17; 24: C*; 25: I*; 26: I16; 27: C20; 28: I26. E' - 29: D12; 30: D14; 31: D11; 32: D9; 33: D10; 34: D13. E" - 35: D20; 36: D29; 37: D32. (F) - 38: D26. * WHO reference strains. Groups are those in Table XXXII (insert). Countries: BR - Brazil; CN - China; CY - Cyprus; ET - Ethiopia; HN - Honduras; IN - India; IR - Iran; IT - Italy; KE - Kenya; LB - Lebanon; PA - Panama; PT - Portugal; SD - Sudan; TU - Tunisia.

Table XXVIII - Jaccard distances within the *L. donovani* complex from pooled RFLP data. Values lower than 0.452 (65% of he maximum distance within the *L. donovani* complex) are shown in black on the upper half.

Group		D							B				A								E									F	C								
	Strain	C	I	C2	I17	I16	C20	I26	D7	D8	D	D4	D3	D2	D15	D21	D30	D27	D25	D16	D9	D11	D14	D12	D10	D13	D20	D29	D32	D26	D1	D18	D19	D5	D17	D24	D33	D28	D34
D	C	-																																					
	I	0.129	-																																				
	C2	0.130	0.130	-																																			
	I17	0.158	0.158	0.092	-																																		
	I16	0.182	0.182	0.130	0.158	-																																	
	C20	0.222	0.222	0.183	0.203	0.222	-																																
	I26	0.294	0.267	0.268	0.282	0.237	0.319	-																															
B	D7	0.622	0.613	0.615	0.619	0.613	0.622	0.621	-																														
	D8	0.616	0.608	0.610	0.613	0.608	0.616	0.616	0.090	-																													
	D	0.608	0.608	0.601	0.605	0.599	0.599	0.616	0.282	0.307	-																												
	D4	0.589	0.589	0.581	0.585	0.579	0.579	0.598	0.267	0.294	0.182	-																											
A	D3	0.554	0.554	0.545	0.550	0.543	0.543	0.565	0.491	0.483	0.437	0.421	-																										
	D2	0.554	0.554	0.545	0.550	0.543	0.543	0.565	0.504	0.496	0.452	0.437	0.130	-																									
	D15	0.579	0.579	0.571	0.575	0.569	0.569	0.589	0.500	0.493	0.491	0.478	0.396	0.396	-																								
	D21	0.587	0.587	0.579	0.583	0.577	0.587	0.587	0.519	0.524	0.535	0.523	0.419	0.419	0.367	-																								
	D30	0.571	0.571	0.573	0.577	0.571	0.581	0.581	0.547	0.541	0.529	0.517	0.397	0.397	0.398	0.386	-																							
	D27	0.592	0.592	0.594	0.598	0.592	0.592	0.601	0.528	0.533	0.532	0.520	0.404	0.404	0.370	0.356	0.352	-																						
	D25	0.577	0.577	0.579	0.583	0.577	0.577	0.587	0.522	0.515	0.514	0.502	0.377	0.377	0.320	0.324	0.273	0.284	-																					
	D16	0.585	0.585	0.587	0.591	0.585	0.585	0.594	0.530	0.523	0.522	0.510	0.388	0.388	0.287	0.291	0.333	0.298	0.229	-																				
E	D9	0.670	0.663	0.665	0.667	0.663	0.663	0.639	0.604	0.607	0.616	0.616	0.637	0.637	0.667	0.661	0.682	0.669	0.671	0.671	-																			
	D11	0.632	0.625	0.627	0.630	0.625	0.625	0.600	0.636	0.632	0.640	0.632	0.637	0.637	0.666	0.680	0.674	0.668	0.677	0.676	0.346	-																	
	D14	0.603	0.594	0.596	0.600	0.594	0.594	0.566	0.608	0.603	0.612	0.603	0.608	0.608	0.642	0.650	0.644	0.637	0.647	0.654	0.388	0.288	-																
	D12	0.636	0.628	0.630	0.633	0.628	0.628	0.603	0.590	0.585	0.594	0.594	0.600	0.600	0.642	0.634	0.651	0.645	0.640	0.647	0.292	0.349	0.385	-															
	D10	0.624	0.624	0.626	0.629	0.616	0.631	0.590	0.643	0.638	0.647	0.647	0.651	0.651	0.697	0.693	0.693	0.674	0.695	0.695	0.460	0.269	0.437	0.437	-														
	D13	0.622	0.614	0.616	0.611	0.606	0.622	0.579	0.642	0.637	0.638	0.638	0.643	0.643	0.691	0.686	0.686	0.667	0.689	0.689	0.468	0.395	0.446	0.446	0.187	-													
	D20	0.583	0.583	0.575	0.579	0.574	0.574	0.564	0.646	0.641	0.626	0.618	0.623	0.623	0.662	0.670	0.664	0.671	0.667	0.673	0.459	0.309	0.390	0.435	0.392	0.369	-												
	D29	0.585	0.585	0.577	0.581	0.575	0.585	0.603	0.616	0.611	0.594	0.594	0.608	0.608	0.649	0.642	0.644	0.652	0.647	0.654	0.473	0.414	0.476	0.450	0.423	0.403	0.293	-										
	D32	0.632	0.632	0.626	0.629	0.632	0.624	0.647	0.644	0.647	0.641	0.641	0.645	0.645	0.688	0.661	0.655	0.663	0.679	0.692	0.488	0.457	0.515	0.492	0.478	0.462	0.392	0.359	-									
F	D26	0.565	0.565	0.567	0.571	0.565	0.555	0.575	0.600	0.594	0.565	0.565	0.551	0.551	0.585	0.583	0.558	0.560	0.544	0.562	0.559	0.542	0.532	0.542	0.559	0.547	0.472	0.461	0.525	-								
C	D1	0.632	0.632	0.634	0.629	0.624	0.624	0.599	0.516	0.510	0.497	0.508	0.558	0.548	0.572	0.606	0.555	0.576	0.570	0.586	0.623	0.623	0.595	0.595	0.622	0.621	0.640	0.634	0.667	0.612	-							
	D18	0.635	0.635	0.645	0.640	0.635	0.643	0.610	0.542	0.536	0.566	0.566	0.608	0.608	0.610	0.641	0.595	0.588	0.600	0.615	0.633	0.633	0.606	0.614	0.610	0.609	0.650	0.658	0.662	0.607	0.393	-						
	D19	0.632	0.632	0.642	0.637	0.632	0.640	0.616	0.538	0.532	0.562	0.562	0.604	0.604	0.615	0.638	0.592	0.593	0.605	0.620	0.631	0.631	0.603	0.611	0.615	0.614	0.654	0.656	0.660	0.612	0.386	0.147	-					
	D5	0.649	0.641	0.651	0.646	0.641	0.649	0.617	0.530	0.524	0.554	0.544	0.597	0.597	0.583	0.631	0.602	0.595	0.590	0.605	0.624	0.632	0.604	0.604	0.631	0.622	0.655	0.670	0.674	0.629	0.421	0.237	0.251	-				
	D17	0.658	0.651	0.660	0.663	0.658	0.658	0.627	0.552	0.547	0.566	0.566	0.608	0.608	0.584	0.583	0.586	0.570	0.565	0.572	0.626	0.654	0.622	0.606	0.667	0.673	0.676	0.678	0.682	0.615	0.463	0.368	0.377	0.349	-			
	D24	0.612	0.612	0.614	0.618	0.612	0.612	0.603	0.530	0.523	0.522	0.522	0.571	0.571	0.575	0.583	0.537	0.570	0.544	0.562	0.643	0.650	0.624	0.624	0.670	0.662	0.622	0.616	0.635	0.562	0.471	0.391	0.415	0.421	0.391	-		
	D33	0.623	0.623	0.625	0.628	0.631	0.614	0.654	0.532	0.536	0.512	0.512	0.573	0.573	0.577	0.594	0.528	0.562	0.566	0.592	0.630	0.651	0.617	0.617	0.678	0.671	0.647	0.609	0.596	0.583	0.497	0.478	0.472	0.500	0.490	0.416	-	
	D28	0.629	0.629	0.631	0.635	0.629	0.629	0.652	0.541	0.545	0.555	0.555	0.600	0.600	0.566	0.583	0.537	0.550	0.554	0.572	0.643	0.663	0.631	0.631	0.683	0.676	0.652	0.624	0.611	0.590	0.518	0.422	0.430	0.449	0.422	0.322	0.300	-
	D34	0.623	0.623	0.625	0.628	0.631	0.623	0.646	0.532	0.536	0.535	0.535	0.583	0.583	0.577	0.604	0.570	0.562	0.575	0.583	0.659	0.678	0.656	0.649	0.685	0.678	0.668	0.656	0.629	0.610	0.530	0.408	0.417	0.407	0.423	0.382	0.383	0.323	-

5.7. Discussion

5.7.1. Targets

A novel PCR-RFLP method using two gp63 intergenic targets (gp63 ITG/L and gp63 ITG/CS) was added to the methods available for phylogenetic analysis of the *L. donovani* complex and phylogeny of the *L. donovani* complex was addressed by a combination of targets for the first time. The four PCR-RFLP analyses complemented each other and proved more robust when pooled. These PCR-RFLPs provided phylogenetic data at different levels, presumably as these gene targets are evolving at different rates.

All targets, except ITG/L, often generated a number of fragments which total length was greater than the size of the amplified fragment. This effect is characteristic of the presence of alleles with different sequences either in the same or in different loci. On the contrary, the total length of ITG/L restriction fragments was mostly smaller than the amplified segment. For some enzymes, as for other targets, this effect was due to not scoring bands larger than 0.9kb. With other enzymes it was apparent that not all possible bands had been scored, probably because some fragments had the same size.

The ITS RFLPs were more conserved but identified genetic diversity between the most divergent strains of *L. donovani*. By ITS RFLP major clusters could be identified by inspection of band profiles.

The gp63 ITG RFLPs distinguished between the *L. infantum / L. chagasi* and *L. donovani* strains. The radical rearrangements in ITG/L of strains D30 (Neal R1) and D31 (Mutinga H9) but also D32 (Jeddah KA) and D33 (Gilani), suggested that ITG/L may suffer different evolutionary rates between groups of strains. ITG/L is multicopy and analysis may be complicated by polymorphisms present in different alleles. However, from analysis of scored total fragment size it emerged that not all fragments may have been differentiated. Furthermore, ITG/L RFLP distances within the *L. donovani* complex were low relatively to ITG/CS and it is possible that ITG/L is very conserved and does not provide enough information. Because of the difficulty in establishing clear groups and the broad inconsistency of the groups by ITG/L as compared with the overall analysis of RFLP data and ITG/CS RFLP, ITG/L emerged as an unreliable marker for typing within the *L. donovani* complex.

The mini-exon RFLPs were also very polymorphic and some groups were not well resolved. The different branching patterns of *L. infantum* by mini-exon RFLP when compared with the other RFLPs might be explained by differential evolution of the numerous copies of the mini-exon unit. The small size (less than 500bp) and the limitation of analysis to four enzymes only, all G and C based, could also have biased

results. The lack of resolution indicated that miniexon RFLP data may not generate reliable markers for strain typing in the *L. donovani* complex, perhaps because of sequence diversity within the mini-exon repeat unit with deletions and repeats (Fernandes *et al.*, 1994). Occasional genetic exchange, in a history of clonal propagation, between *Leishmania* parasites could complicate phylogenetic analysis. Topology of the UPGMA tree was affected by assumption of a molecular clock, as seen by comparison with the neighbour-joining tree and the Fitch-Margoliash tree. Position of strain D29 (Addis 164) changed radically, but other groups were also badly affected. *Leishmania infantum* / *L. chagasi*, Indian *L. donovani* and group E (ILM 10) were the most reliable. Both because phylogenies are not congruent and because of the low bootstrap values, mini-exon does not seem to be useful on its own for phylogenetic analysis of the *L. donovani* complex. Despite this, some mini-exon RFLP markers may be useful for typing, although its use will be limited.

Compared with the pooled analysis of RFLP data, the best targets were ITG/CS, followed by ITS, although with less resolution, and the least useful targets were ITG/L and mini-exon, which did not seem to generate reliable markers for *L. donovani* strain typing. The latter targets were also those which generated the smallest number of characters and the least clear cut markers which may have influenced the analyses.

The general consistency of the groups of strains upon phylogenetic analyses indicates the suitability of RFLP of intergenic regions for taxonomic studies. Affinities between groups varied depending on the region analysed, which could indicate different rates of evolution, a radiation event or be due to limitions of the methods of analysis. It is probable, however, that individual targets simply do not produce enough markers to produce reliable trees because most individual analyses had low bootstrap support throughout, whilst the pooled analysis was robust.

5.7.2. Phylogenetic methods

Both Wagner and Dollo parsimony methods performed well, regarding tree lengths and similarity to distance methods, and the main branches in the resulting trees for the pooled RFLP analysis had high bootstrap support. These methods also generated the shortest trees, when compared with Camin-Sokal and polymorphism parsimony methods. This may be because the first allows reversibility of characters, whilst Camin-Sokal does not. The robustness of the trees, measured by bootstrap analysis, seemed to decrease with the length of the trees generated, which agrees with the theoretical basis of cladistics. Wagner parsimony produced the shortest and most robust trees and may be, in the future, the parsimony method of choice to

handle RFLP analysis of large data sets, despite that Dollo is the best theoretical method.

Distance methods seemed to perform well with RFLP data. All tested methods generated the same major genetic groups, despite some differences in topology. This effect may be due to independence of a molecular clock on the part of both Fitch-Margoliash and neighbour-joining methods. Indeed the cophenetic correlation coefficient of the UPGMA tree was 0.97, which indicated that the data were almost but not perfectly ultrametric.

Phenetic (distance) and cladistic (parsimony) methods produced similar results, particularly when the largest, pooled, data set was used. Parsimony methods may in these instances behave as phenetic methods and thus with pooled RFLP data true cladograms of the *L. donovani* complex may not have been obtained.

5.7.3. Taxa within the *Leishmania donovani* complex

In all PCR-RFLP analyses *L. chagasi* was indistinguishable from *L. infantum* as had been described previously (Beverley *et al.*, 1987; Cupolillo *et al.*, 1994; Moreno *et al.*, 1984; Schonian *et al.*, 1996; van Eys *et al.*, 1991) and, upon phylogenetic analysis, *L. infantum* (*L. chagasi*) were monophyletic except by mini-exon RFLP. The degree of diversity found here fell within the level of intraspecific diversity (Beverley *et al.*, 1987) and each *L. donovani* group had similar or more genetic diversity than *L. infantum* (*L. chagasi*).

Four more strain clusters were identified within the *L. donovani* complex in a pooled analysis of all RFLPs. Those groups had also been found in most individual analyses, although some were not well resolved, especially by ITG/L and mini-exon RFLPs. It was apparent that the designation *L. donovani* includes at least four main genotypic groups, equivalent to *L. infantum*. When trees were rooted (ITS and mini-exon RFLPs, or by using UPGMA), the most ancient branch of the *L. donovani* complex was more often with the cluster identified as group E (strains from Ethiopia and Saudi Arabia) in the pooled analysis, although *L. infantum* was often not far away from the root, thus challenging alternative rootings for the *L. donovani* complex (Rioux *et al.*, 1990).

6. Comparison of DNA Sequences. The Major Surface Protein Genes

6.1. Introduction

6.1.1. DNA sequencing

Sequencing of a DNA molecule can be achieved through two major lines of methodology, developed in the late seventies. The Maxam and Gilbert, or chemical method, is founded on base-specific or base-selective cleavage of DNA (Maxam and Gilbert, 1977; Maxam and Gilbert, 1980). The Sanger, or dideoxy method, is based on addition of complementary bases to a template DNA strand (Sanger *et al.*, 1977). Both methods required separation by polyacrylamide gel electrophoresis of labeled DNA fragments, which was improved with the development of thin polyacrylamide gels. The dideoxy method, the most widely used, is more flexible and requires less labour than chemical methods which, however, are irreplaceable in locating rare bases, such as 5-methylcytosine sites.

In the Sanger method a complementary strand is synthesized from single stranded DNA template, with the help of a priming oligonucleotide. Introduction of di-deoxynucleotides (ddNTPs) interrupts synthesis of the complementary strand, thus producing different sized DNA fragments. Either the primer or terminator ddNTPs can be labeled in order to identify the end base of each DNA fragment. Radioactive labeling, always requires four reactions for identification of the four bases, whilst with fluorescent labeling, differently labeled terminators can be used in one tube reactions. Fluorescent labeling also allows fragment separation in one lane, whichever labeling method is used.

Because both sequencing methods require single stranded DNA, cloning of template DNA in single stranded vectors such as the bacteriophage M13 is desirable. By adaptation of the polymerase chain reaction (PCR) to the dideoxy method, sequencing of double stranded DNA was made possible in what is now called direct sequencing. This method involves less manipulations of template DNA, is faster and, with adequate primers, sequences of both DNA strands are easily obtained. PCR products can also be sequenced. Furthermore, fluorescently labeled products can be separated in automated sequencers, which detect the passage of labeled DNA fragments through a point in the gel. Sequencing with fluorescent labeling is faster than with radioactive labeling, not only because the detection step is simultaneous with electrophoresis, but readable data are immediately produced. It is also safer, but less sensitive.

In the present work, direct sequencing of PCR products was done with fluorescent labeling. The same primers as for the preparative PCR reaction were often used. Primers for sequencing are designed according to the same rules as PCR primers, although more demanding in terms of specificity, because non-specific products, can make a sequencing reaction unreadable.

6.1.2. Major surface protease genes

Major surface protease (*msp*) or GP63 coding genes are important antigenic genes in *Leishmania* (Morales *et al.*, 1997; Russo *et al.*, 1991; Shreffler *et al.*, 1993; Yang *et al.*, 1993). These genes seem to be important for parasite virulence and may have some role in internalization of the promastigote in the macrophage (Chakrabarty *et al.*, 1996). Their proteolytic activity has been implicated in cleavage of CD4 which may influence induction of the immune response and thus disease progression in *Leishmania* infections (Hey *et al.*, 1994). GP63 may (Brittingham *et al.*, 1995; Russell, 1987) or may not (Nunes and Ramalho Pinto, 1996) be important for complement fixation, but seems to protect against complement mediated lysis (Joshi *et al.*, 1998). Analysis of *msp* sequence may provide relevant comparisons with current isoenzyme classification or information on adaptation to hosts and vectors, and a more functional perspective than the intergenic regions analyzed by RFLP (Chapter 5).

Msp genes have been found in multiple copies in all *Leishmania* species for which they have been studied (Button *et al.*, 1989; Maingon *et al.*, 1990; Medina Acosta *et al.*, 1993; Ramamoorthy *et al.*, 1992; Steinkraus *et al.*, 1993; Victoir *et al.*, 1995). Three gene families were described in *L. chagasi* (Ramamoorthy et al., 1992): *mspL*, *mspS* and *mspC*. *MspC* was identified as being a single copy gene, while the other gene families were present in multicopy arrays (Roberts *et al.*, 1993; Webb *et al.*, 1991). Multigenic gene clusters have also been identified in *L. mexicana* (Medina Acosta *et al.*, 1993) where a gene cluster of *mspC* type genes was described (gp63 C1). *MspC* genes do not have consensus GPI anchors and the enzyme product has been located in the flagellar pocketof the amastigote (Medina Acosta *et al.*, 1993). The homologous gene isolated from *L. major* (gp63-C) was similarly expressed in the amastigote, but the carboxyl terminus was found to be capable of mediating GPI anchor attachment (Voth *et al.*, 1998). *Leishmania major* also has different classes of GP63 (Button *et al.*, 1989; Voth *et al.*, 1998), which are differentially expressed during the parasite cycle. Although different *Leishmania* species seem to have distinct GP63 gene organizations and copy numbers, the presence of different gene classes in cluster organization seems to be a rule.

Phylogenetic reconstruction based on multicopy genes is fraught with methodological problems. Phylogenetic trees may not reflect true genealogies but, instead, composite genealogies based on different loci. Thus, for simplification of approach, the single gene *mspC* was chosen for sequencing and analysis.

Leishmania chagasi mspC has two regions involved in the Zinc binding site (Ramamoorthy *et al.*, 1992), but not a recognizable GPI anchor binding site. It is debatable whether or not this gene codes for a soluble enzyme (Medina Acosta *et al.*, 1989). *MspC* is unique among other *msp* genes in having a longer and distinct 3' or carboxyl terminus, which may encode for a transmembrane region or for mediation of GPI attachment (Voth *et al.*, 1998), at least in *L. major*. Specific sequences were identified which enabled design of primers for selective PCR amplification of the 3' end of this gene (*mspC3*).

6.2. PCR amplification of *mspC*

Partial *mspC* amplification was achieved as described in Materials and Methods (Chapter 2). However, the thermal profile was later modified to increase the yield whilst maintaining specificity of PCRs to amplify the products B and C (Materials and Methods). Previous conditions were kept, except that the stage with annealing temperature at 65°C was reduced to 5 cycles and followed by 25 cycles with an annealing temperature of 60°C.

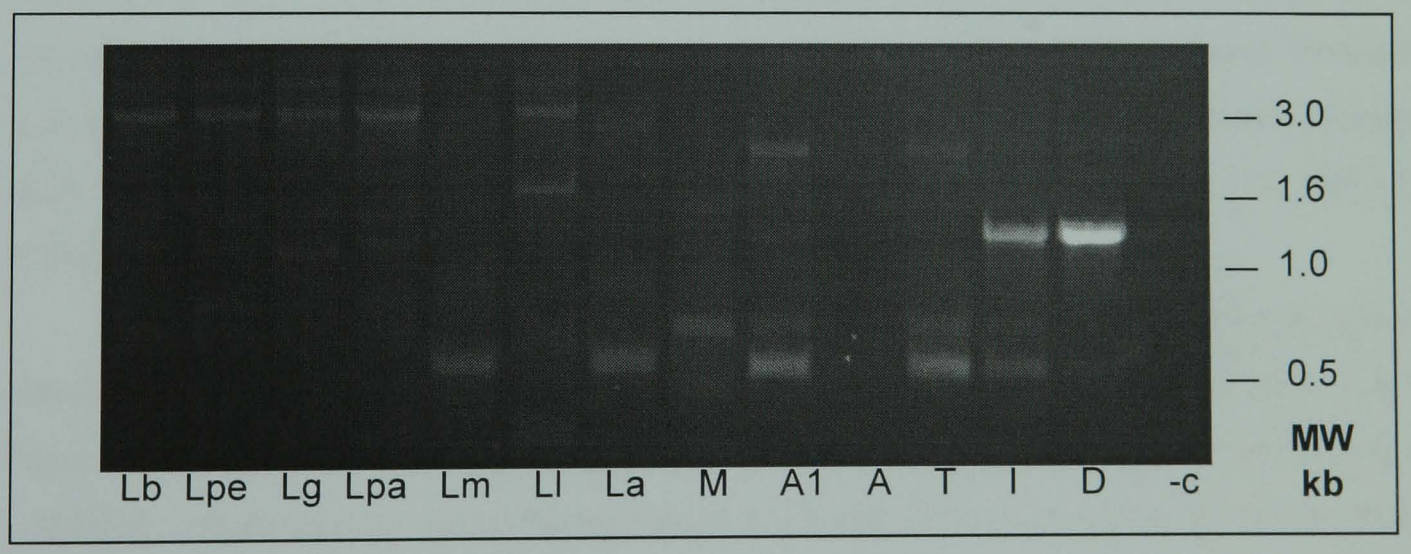


Figure 38 - Specificity of *mspC* amplification for the *L. donovani* complex. -c is negative control; Lb is *L. braziliensis*, Lpe is *L. peruviana*, Lg is *L. guyanensis*, Lpa is *L. panamensis*; Lm is *L. mexicana*, Ll is *L. lainsoni*, La is *L. amazonensis*, M is *L. major*, A is *L. aethiopica* (L100) and A1 is L96, T is *L. tropica*, I is *L. infantum* and D is *L. donovani*.

Amplification of each fragment or the entire *mspC3* was specific for *L. donovani* complex strains and generated a single product of 1.3-1.4 kb (Fig. 38). Amplification with strain Sukkar 2 (D23) was only achieved for the C fragment, which generated two different sized products of around 500bp (Fig. 39).

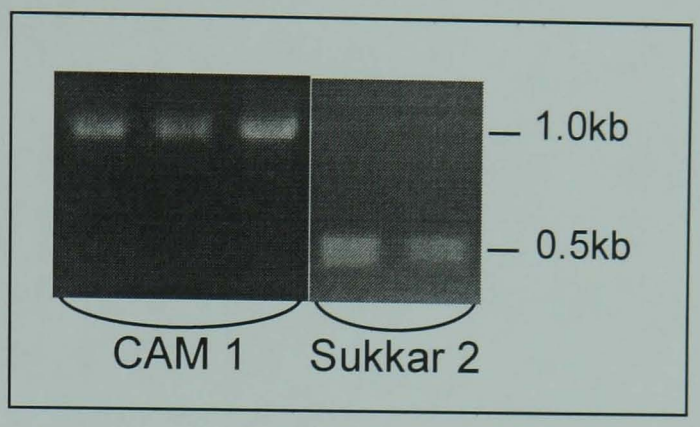


Figure 39 - Double band *mspC* - C product generated for strain Sukkar 2 (D23) compared with a full *mspC* product (CAM 1, I33).

6.3. *Leishmania infantum* and *Leishmania chagasi* are indistinguishable

The full 1083bp sequences from the coding region of *L. infantum* and *L. chagasi mspC3* were identical (strains IPT-1, PP75, DOG124, M6445, Strain A, IMT171, Buck; EMBL accession numbers AJ010234, AJ009908, AJ009911, AJ010240-2, AJ290785 respectively), except for a single base pair in strain WR285 (at nucleotide 582; EMBL accession number AJ009909). All incomplete *mspC3* sequences from other *L. infantum* and *L. chagasi* strains were also identical to the complete sets: combined sequences of fragments A and B for strains Pharoah, Lombardi, LEM75, Alessandro (599-625 nucleotides (nt), EMBL accession numbers AJ010248-51), and of fragments B and C for L82, Rebelo 2, L53, IMT170, IMT172, CO910 (784-840nt, EMBL accession numbers AJ010243-47, AJ009910).

All sequences (including those from *L. donovani* strains) possessed highest similarity to *L. chagasi mspC* on a BLAST search (Experimental WU-BLAST server from the Bioinformatics Group of the Swiss Institute for Experimental Cancer Research - ISREC), confirmed by visual inspection, but diverged from it by three extra base pairs in the coding region (Fig. 40).

```
A L. chagasi (published) 5'catcccttggg--ctactcgc-atttct3'
  L. donovani complex    5'catcccttggggctactcgccatttct3'

B  L. chagasi (published)     S  L  G    L  L  A   F
   L. donovani complex        S  L  G  G  Y  S  P  F
```

Figure 40 - Divergence of *L. donovani* complex *mspC* sequences from the originally published *L. chagasi mspC* (M80671, nucleotides 1450 to 1474). A is nucleotide sequence and B is amino acid sequence (centre of codon). Dashes (-) were inserted in the sequence to maximize alignment.

6.4. Genetic variability among the *Leishmania donovani*

Full *mspC3* protein coding sequences were obtained from the *L. donovani* strains listed below. The groups listed below could be formed based on the degree of identity between sequences. Only one sequence of which was used for phylogenetic reconstruction:

A. A1. MRC(L)3 (D2, EMBL AJ010238) and Mutinga H9 (D31, EMBL AJ290780);
   A2. MRC74 (D3, EMBL AJ010239) ;
   A3. Ndandu 4A (D16, EMBL AJ290770)
   A4. LRC-L57 (D21, EMBL AJ290774) and D2 (D25, EMBL AJ290776);
   A5. Ayele 8 (D27, EMBL AJ290777);
B. Patna 1 (D4, EMBL AJ010236), Chowd-X (D6, EMBL AJ290783) and STL1-79 (D7, EMBL AJ290784); which were very similar to DD8 (D, EMBL AJ010235);
C. HU3 (D1, EMBL AJ010237), Salti 4 (D17, EMBL AJ290771), Khartoum (D18, EMBL AJ290772), A22 (D19, EMBL AJ290773), Dora (D24, EMBL AJ290775), Gebre 1 (D28, EMBL AJ290778), Gilani (D33, EMBL AJ290781), IMT180 (D34, EMBL AJ290782) and IMT 188 (D35, EMBL AJ290786);
D. *L. infantum* and *L. chagasi* strains (except WR285);
E. VL29 (D10, GenBank AF267730), Ayele 5 (D12, EMBL AJ290768), Hussen (D13, EMBL AJ290769) and Addis 164 (D29, EMBL AJ290779);

Both Sukkar 2 (D23) *mspC3* - fragment C PCR products were sequenced and an extensive deletion of 48bp (Fig. 41) was found in the shorter fragment. The sequence corresponding to the gap was flanked on both sides by 'gcggcg' which in the shorter fragment was reduced to one 'gcg'. Otherwise both sequences were identical. The stop codon present in all other strains was also modified to code for

serine, but there was a second stop codon 11 amino acids dowstream (Fig. 43 and 45).

Fifty six nucleotide positions (5.2%) were found to be polymorphic within the studied strains of the *L. donovani* complex (Sukkar 2 excluded), which corresponded to 28 amino acid polymorphisms (7.78%) (Fig. 43), including changes in the second region of the Zinc binding site in the *L. donovani* strains (Ramamoorthy *et al.*, 1992). Eighteen polymorphisms were observed in the first codon position, 15 in the second and 24 in the third. All substitutions in first and second positions were non-silent, but third position substitutions were all silent except one. One of the third position polymorphisms was acompanied by a second position polymorphism, which meant that a third position substitution could have been silent or not, depending on whether the adjacent second position substitution occurred before or after it. Although there were 33 non-silent substitutions, in some cases there was more than one non synonymous polymorphic site per codon (three codons with two and one codon with three), and thus only 28 amino acid polymorphisms.

Despite the large number of aminoacid polymorphisms, most substitutions were unlikely to modify the function of the protein. A large number of putative aminoacid substitutions were for aminoacids with similar net charge (31.0%): 5 (17.2%) polar and 4 (13.8%) hydrophobic. No changes between negative and positively charged aminoacids were observed, but 34.5% were between polar and negative (5) or positively (5) charged residues. However, a large number, in a total of 37.9%, were changes from hydrophobic to polar (6) or charged aminoacids (5). These polymorphisms may be responsible for different properties of the proteins in which they are present. Most changes among the six observed between *L. infantum* and group E (D10) were between polar and hydrophobic amino acids (four), whilst the remainder were between polar and charged residues.

A number of polymorphisms might be considered specific for *L. infantum* / *L. chagasi* and group E in relation to the majority of the *L. donovani* (Fig. 42). However, some of these were not specific in relation to the *L. major* and *L. mexicana* sequences. Within the *L. donovani* complex position 932 is specific for group E, and positions 1011 and 1020 for *L. infantum* (*L. chagasi*) .

```
Group                    972(M80671)                                                                  bp
D   L. infantum      caaatgatga acattcgtgg aaaggacttt aatgtttctg tgatcaacag cagcacggcg gtggcgaagg 70
E   L. donovani 10   .......... .......... .........c g.....c.c. .......... .......... ..........
B1  L. donovani      ..g....... .......... .......... .......... .......... .......... ..........
B   L. donovani 4    .......... .......... .......... .......... .......... .......... ..........
C   L. donovani 1    .......... .......... .......... ......c.c. .......... .......... ..........
A1  L. donovani 2    .......... .......... .......... .......... .......... .......... ..........
A2  L. donovani 3    .......... .......... .......... .......... .......... .......... ..........
A3  L. donovani 16   .......... .......... .......... .......... .......... .......... ..........
A4  L. donovani 21   .......... .......... .......... .......... .......... .......... ..........
A5  L. donovani 27   .......... .......... .......... .......... .....c.... .......... ..........
    L. major         a.cg.ttc.. ..g....a.. c...a....c g.....c.c. .......... .........a ..........
    L. mexicana      g.gg..cc.c ..c....gc. c......... .....g..g. ......c... ........t. ..........

D   inf cgcgcgagca gtacggctgc gacaccttgg agtatctgga gatcgaggac cagggcggtg cgggctccgc cgggtcgcac 150
E   D10 .......... .......... .......... .......... .......... .......... .......... ..........
B1  D   .......... .......... .......... .......... .......... .......... .......... ..........
B   D4  .......... .......... .......... .......... .......... .......... .......... ..........
C   D1  .......... .......... .g........ .......... .......... .......... .a........ ..........
A1  D2  .......... .......... .......... .......... .......... .......... .......... ..........
A2  D3  .......... .......... .......... .......... .......... .......... .......... ..........
A3  D16 .......... ......g... .......... .......... .......... .......... .......... ..........
A4  D21 .......... .......... .......... .......... .......... .......... .......... ..........
A5  D27 .......... .......... .......... .......... .......... .......... .......... ..........
    maj .......... .......... .......... .......... .g.g...... .......... .......... ..........
    mex ....t..... .......... a...g..... .......... ...t...... .......... .......... .........t
```

(Figure 41 - Alignment of *L. donovani* complex and outgroup *mspC3* sequences. See page 161)

Group									bp
D	inf	atcaagatgc	gcaacgccaa	ggacgagctc	atggcgcctg	ccgcagctgc	cgggtactac	agcgccctga	ccatggccat	230
E	D10	.....	.....	.....	.....	.....	.....	.....	.....	
B1	D	.....	.....	.....	.....	.....	.....	.....	.....	
B	D4	.....	.....	.....	.....	.....	.....	.....	.....	
C	D1	.....	.....	.....	.....	.....	.....	.....	.....	
A1	D2	.....	.....gc.	.....	.....	.....	.....	.....	.....	
A2	D3	.....	.....gc.	.....	.....	.....	.....	.....	.....	
A3	D16	.....	.....	.....	.....	.....	.....	.....	.....	
A4	D21	.....	.....	.....	.....	.....	.....	.....	.....	
A5	D27	.....	.....	.....	.....	.....	.....	.....	.....	
	maj	.....	.....gc.	.....	.....	.t.....	.....	.c.....	.....	
	mex	.....	.....	.....	.....	....t....	.....	.c.....	.....g.	

D	inf	cttccaggac	ctcggcttct	accaggcgga	cttcagcaag	gccgaggtga	tgccgtgggg	ccggaacgcc	ggctgcgcct	310
E	D10	.....	.....	.....	.....	.....	.....	.....	.....	
B1	D	.....	.....	.....	.....	.....a..	.....	.....	.....	
B	D4	.....	.....	.....	.....	.....a..	.....	.....	.....	
C	D1	.....	.....	.....	.....	.....a..	.....	.....	.....	
A1	D2	.....	.....	.....	.....	.....a..	.....	.....	.....	
A2	D3	.....	.....	.....	.....	.....	.....	.....	.....	
A3	D16	.....	.....	.....	.....	.....a..	.....	.....	.....	
A4	D21	.....	.....	.....	.....	.....a..	.....	.....	.....	
A5	D27	.....	.....	.....	.....	.....a..	.....	.....	.....	
	maj	.c.....	.....	.....	.....	.....	.....	..a.....	.....	
	mex	.....	.....	.....	.....	.....a..	.....	.....t.	.....	

(Figure 41 - Alignment of *L. donovani* complex and outgroup *mspC3* sequences. See page 161)

```
Group                                                                                                                  bp
D  inf  tcctcagcga gaagtgcatg gagcggaaca tcacgaagtg gccggcgatg ttctgcaatg agaacgaggt gactatgcgc                        390
E  D10  .......... .......... .......... .......... .......... .......... .......... ..........
B1 D    .......... .......... ....a..... .......... .......... .......... t..gt.t..a cgt.g.....
B  D4   .......... .......... ....a..... .......... .......... .......... t..gt.t..a cgt.g.....
C  D1   .......... .......... ....a..... .......... .......... ..t....... t..gt.t..a cgtcg.....
A1 D2   .......... .......... ....a..... .......... .......... ..t....... t..gt.t..a cgt.g.....
A2 D3   .......... .......... ...gacgg.. .......... .......... .......... t..gt.t..a cgt.g.....
A3 D16  .......... .......... ....a..... .......... .......... ..t....... t..gt.t..a cgt.g.....
A4 D21  .......... .......... ....a..... .......... .......... ..t....... t..gt.t..a cgt.g.....
A5 D27  .......... .......... ....a..... .......... .......... ..t....... t..gt.t..a cgt.g.....
   maj  ......c.a. c......... ....a..g.g .....c.... .......... .......... ...g.....a cg.c..c...
   mex  .......... .......... .c.aa....g .......... .......... .......... ...gt.c..c c..c..a..g

D  inf  tgccccacca gtcgtctgat ggtcggaacc tgtggtataa ggggatacag cactccgttt tcgctgtact ggcagtactt                        470
E  D10  .......... .......... .......... .......... .......... .......... .......... ..........
B1 D    .......... .......... .c........ .......... .a........ .......... ....cc.... ..........
B  D4   .......... .......... .c........ .......... .a........ .......... ....cc.... ..........
C  D1   .......... .......... .c........ .......... .a........ .......... ....cc.... ..........
A1 D2   .......... .......... .c........ .......... .a........ .......... ....c..... ..........
A2 D3   .......... .......... .c........ .......... .a........ .......... ....c..... ..........
A3 D16  .......... .......... .c........ .......... .a........ .......... ....c..... ..........
A4 D21  .......... .......... .c........ .......... .a........ .......... ....c..... ..........
A5 D27  .......... .......... .c........ .......... .a........ .......... ....c..... ..........
   maj  .......... .......cc. cc.......g .......... ...a....ga ac.g....g c...ga.... ..........
   mex  .........g ac.......g a........t .......... ca.c.....a t...t....g g..ac..... ..........
```

(Figure 41 - Alignment of *L. donovani* complex and outgroup *mspC3* sequences. See page 161)

```
Group                                                                                                  bp
D  inf  caccaacgca tcccttgggg gctactcgcc atttctggat tactgcccgt ttgttatcgg ctacagtgat ggttcgtgca 550
E  D10  .......... .......... .......... .......... .......... .......... .......... .......... 
B1 D    ......tat. .......... .......... .......... .......... .......... ....g..... .......... 
B  D4   ......tat. .......... .......... .......... .......... .......... ....g..... .......... 
C  D1   ......tat. .......... .......... .......... .......... .......... ....g...g. .......... 
A1 D2   ......tat. .......... .......... .......... ..t....... .......... ....g..... .......... 
A2 D3   ......tat. .......... .......... .......... ..t....... .......... ....g..... .......... 
A3 D16  ......tat. .......... .......... .......... ..t....... .......... ....g..... .......... 
A4 D21  ......tat. .......... .......... .......... ..t....... .......... ....g..... .......... 
A5 D27  ......tat. .......... .......... .......... ..t....... .......... ....g..... .......... 
   maj  .........g ..t....... .......... ...c.....c .......... .......... ....gc.... ..c.......
   mex  ......t..g .....c.... .......... ...c.....c .......... .....g.t.. ......ga.. ..c.......

D  inf  atcaggacgc atcgttggca gcagggtttt tcagtgcatt caacgtcttc tccgacgcgg cgcgctgcat cgatggcgcc 630
E  D10  .......... .......... .......... .......... .......... .......... .......... .......... 
B1 D    .......... .......... a......... ..g....... .......... .......... .......... .......... 
B  D4   .......... .......... a......... ..g....... .......... .......... .......... .......... 
C  D1   .......... .......... a......... ..g....... .......... .......... .......... .......... 
A1 D2   .......... .......... a......... ..g....... .......... .......... .a........ .......... 
A2 D3   .......... .......... a......... ..g....... .......... .......... .a........ .......... 
A3 D16  .......... .......... a......... ..g....... .......... .......... .a........ .......... 
A4 D21  .......... .......... a......... ..g....... .......... .......... .a........ .......... 
A5 D27  .......... .......... a......... ..g....... .......... .......... .a........ .......... 
   maj  .......... .....c.... .ag.a...c. ...c...... .......... ..t....... .......... .......... 
   mex  .......t.. g...ac.a.. c.g.acc..c ..gc...g.. .......... .....g..c. ....g..... .......... 
```

(Figure 41 - Alignment of *L. donovani* complex and outgroup *mspC3* sequences. See page 161)

```
Group                                                                                                  bp
D  inf  ttcaggccga agaatagaac cgctgccaat ggctactacg ccggactgtg cgccaacgtg cggtgcgaca cggccacgcg 710
E  D10  ...c...... .......... .......... .......... .......... .......... .......... .......... 
B1 D    .......... .......... .......g.. .......... .......... .......... .......... .......... 
B  D4   .......... .......... .......g.. .......... .......... .......... .......... .......... 
C  D1   .......... .......... .......g.. .......... .......... .......... .......... .......... 
A1 D2   .......... .......... .......g.. .......... .......... .......... .......... .......... 
A2 D3   .......... .......... .......g.. .......... .......... .......... .......... .......... 
A3 D16  .......... .......... .......g.. .......... .......... .......... .......... .......... 
A4 D21  .......... .......... .......g.. .......... .......... .......... .......... .......... 
A5 D27  .......... .......... .......g.. .......... .......... .......... .......... .......... 
   maj  .....a.... ..gca.ct.a ..gcatagtc aag.cg.... ....c..... .......... .a...t.... .......a.. 
   mex  ....c..... ....c..... ......gg.. ..a......a ...cc..... .......... aa........ .......... 

D  inf  cacgtacagc gtgcaggtgc gcggcagtat ggactacgtg aactgcacgc cgggcctcag agttgagctg agcaccgtga 790
E  D10  .......... .......... .......... .......... .......... .......... .......... .......... 
B1 D    .......... .........t .......... .......... ........c. .......... .......... .......... 
B  D4   .......... .........t .......... .......... ........c. .......... .......... .......... 
C  D1   .......... .........t .......... .......... ........c. .......... .......... ...t...... 
A1 D2   .......... .......... .......... .......... ........c. .......... .......... .......... 
A2 D3   .......... .......... .......... .......... ........c. .......... .......... .......... 
A3 D16  .......... .......... .......... .......... ........c. .......... .......... .......... 
A4 D21  .......... .......... .......... .......... ........c. .......... .......... .......... 
A5 D27  .......... .......... .......... .......... ........c. .......... .......... .......... 
   maj  .......... .......... a........a c......acc .......... .......... .......... .......... 
   mex  .......... ..c....... ......cg.a c.g....cc .......... .......... ....a..t.. ....g..... 
   D23a                                  .... .g..... .......... .......... .......... .......... 44
   D23b                       .......... .......... .g..... .......... .......... .......... .......... 60
```

(Figure 41 - Alignment of *L. donovani* complex and outgroup *mspC3* sequences. See page 161)

```
Group                                                                                                  bp
D   inf   gcagcgcctt cgaggagggc ggctacatca cgtgcccgcc gtacgtggag gtgtgccagg ccaacgtcaa gggagccaag 870
E   D10   .......... .......... .......... .......... .......... .......... g......gc. ..ct......
B1  D     .......... .......... .......... .......... .......... .......... .......g.. ..........
B   d4    .......... .......... .......... .......... .......... .......... .......g.. ..........
C   d1    .......... .......... .......... .......... .......... .......... .......g.. ..........
A1  d2    .......... .......... .......... .......... .......... .......... .......... ..........
A2  d3    .......... .......... .......... .......... .......... .......... .......... ..........
A3  d16   .......... .......... .......... .......... .......... .......... .......... ..........
A4  d21   .......... .......... .......... .......... .......... .......... .......... ..........
A5  d27   .......... .......... .......... .......... .......... .......... .......... ..........
    maj   ...aaa.... ......a... .......... .......... .......... .......... g......gc. ..ct......
    mex   ..ga...... ....a..... ......g... .......... .......... .......... g......... a.ct......
    D23a  ...a...... .......... .......... .......... .......... .......... g......... .......... 124
    D23b  ...a...... .......... ....n..... .......... .......... .......... g......... .......... 140

D   inf   gacttcgcag gcgactccga cagctccagc agcgccggtg acgctgccga cagagcggcg atgcagcggt ggaatgacag 950
E   D10   .......... .......... .......... .......... .......... .......... .c........ ..........
B1  D     .......... .......... .......... .......... .......t.. .......... .......... ..........
B   D4    .......... .......... .......... .......... .......t.. .......... .......... ..........
C   D1    .......... .......... .......... .......... .......t.. .......... .......... ..........
A1  D2    .......... .......... .......... .......... .......t.. .......... .......... ..........
A2  D3    .......... .......... .......... .......... .......t.. .......... .......... ..........
A3  D16   .......... .......... .......... .......... .......t.. .......... .......... ..........
A4  D21   .......... .......... .......... .......... .......t.. .......... .......... ..........
A5  D27   .......... .......... .......... .......... .......t.. .......... .......... ..........
    maj   .....t.at. .......... .......... ...t..a... .......... ..ag...... ..tg...... .......g..
    mex   .....t.... .....a.... .......... .......a.. .......... ..a..a.... .......... ...g......
    D23a  .....t..t. .......... .......... .......a.. .........g ..ag...... .......... .......... 204
    D23b  .....t..t. .......... .......... .......a.. .........g ..ag...--- ---------- ---------- 197
```

(Figure 41 - Alignment of *L. donovani* complex and outgroup *mspC3* sequences. See page 161)

Group									bp
D	inf	gatggccggc	ttggctactg	cggcgatggt	gctgctagga	atggttctct	ctctcatggc	actcgtggtg	1020
E	D10	....	....	....	....	....	....	g.......t	
B1	D	....	....	....	....	....	....	g.......t	
B	D4	....	....	....	....	....	....	g.......t	
C	D1	....	....	....	....	....	....	g.......t	
A1	D2	....	....	....	....	....	....	g.......t	
A2	D3	....	....	....	....	....	....	g.......t	
A3	D16	....	....	....	....	....	....	g.......t	
A4	D21	....	....	....	....	....	....	g.......t	
A5	D27	....	....	....	....	....	....	g.......t	
	maj	....	c.........	......c...	....	g.a.......	....	......c...	
	mex	........c.	....	..a...c.c.	....	.....g....	....	....c.c...	
	D23a	....	....	......c...	....	....	....	......c...	274
	D23b	----------	----------	-----.c...	....	....	....	......c...	242

D	inf	gtgtggctac	tccttctcac	ctgcccctgg	tggtgttgca	aatttggggg	gctcccgacg	tga	1083bp
E	D10	....	....	....	....	....	....	...	
B1	D	....	....	....	....	....	......a...	...	
B	D4	....	....	....	....	....	......a...	...	
C	D1	....	....	....	....	....	......a...	...	
A1	D2	....	....	....	....	....	......a...	...	
A2	D3	....	....	....	....	....	......a...	...	
A3	D16	....	....	....	....	....	......a...	...	
A4	D21	....	....	....	....	....	......a...	...	
A5	D27	....	....	....	....	....	......a...	...	
	maj	....	.g...g...g	.......a..	....	..g.......	....	...	
	mex	...c......	.....ac..g	..c.......	..c..c....	g.c.g.....	....	...	
	D23a	....	.....g...g	....	....	..c.......	..ct...g..	.c.	337bp
	D23b	....	.....g...g	....	....	..c.......	..ct...g..	.c.	305bp

Figure 41 - Alignment of *L. donovani* complex and outgroup (*L. major* - maj - and *L. mexicana* - mex) *mspC3* sequences. Dots are identical bases and (-) are missing bases. Double underlined are stop codons. Lightly shaded boxes are nucleotides specific to *L. donovani* in relation to *L. infantum / L. chagasi* and group E (D10) and dark boxes are nucleotides specific to all *L. donovani* in relation to all *L. infantum / L. chagasi*. Groups are the same as those in Table XXXII (insert).

Groups		335	371	385	412	432	455	477	535	571	583	634	658	759	918	932	1011	1020	1077
D	*L. infantum*	g	agaacgaggtgacta		g	g	t	cgc	a	g	ca	a	a	g	c	t	a	g g	
E	VL29	.	.....		.	.	.	...	.	.	..	c	.	.	.	c	g	t .	
B1	DD8	a	t..gt.t..acgt.g		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
B	Patna 1	a	t..gt.t..acgt.g		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
C	HU3	a	t..gt.t..acgtcg		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
A2	MRC74	a	t..gt.t..acgt.g		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
A1	MRC(L)3	a	t..gt.t..acgt.g		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
A3	Ndandu 4A	a	t..gt.t..acgt.g		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
A5	Ayele 8	a	t..gt.t..acgt.g		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
A4	LRC-L51	a	t..gt.t..acgt.g		c	a	c	tat	g	a	.g	.	g	c	t	.	g	t a	
	L. major	a	...gc....acgcc.		c	.	g	...	g	.	..	.	g	.	.	.	.	. .	
	L. mexicana	a	...gt.c..cc..c.		.	a	c	t..	.	c	.g	.	g	.	.	.	.	. .	

Figure 42 - Sites potentially diagnostic between *L. infantum* (*L. chagasi*), group E strains (VL29, Ayele 5, Hussen and Addis 164), and *L. donovani*. Underlined are positions with sequence specific to *L. infantum* within the *L. donovani* complex and double underlined are positions specific to group E among the species studied. Nucleotide positions as in Fig. 41. Groups are the same as those in Table XXXII (insert).

```
Group                                  ×          +   ✳                                                                  aa
D     L. infantum      QMMNIRGKDF NVSVINSSTA VAKAREQYGC DTLEYLEIED QGGAGSAGSH IKMRNAKDEL MAPAAAAGYY 70
E     L. donovani 10   .......... D.P....... .......... .......... .......... .......... .......... ..........
B=B1  L. donovani 4    .......... .......... .......... .......... .......... .......... .......... ..........
C     L. donovani 1    .......... ..P....... .......... G......... .......... .......... .......... ..........
A1    L. donovani 2    .......... .......... .......... .......... .......... .......... ......Q... ..........
A2    L. donovani 3    .......... .......... .......... .......... .......... .......... ......Q... ..........
A3    L. donovani 16   .......... .......... .......... .......... .......... .......... .......... ..........
A4    L. donovani 21   .......... .......... .......... .......... .......... .......... .......... ..........
A5    L. donovani 27   .......... .....H.... .......... .......... .......... .......... .......... ..........
      L. major         NVS.V...N. D.P....... .......... .......V.. .......... .......... ......Q... ..........
      L. mexicana      EVPHL.R... .....T...V .......... NS........ .......... .......... .....S....

                                                                  +                  zn
                                                                                     ×
D     INF SALTMAIFQD LGFYQADFSK AEVMPWGRNA GCAFLSEKCM ERNITKWPAM FCNENEVTMR CPTSRLMVGT CGIRGYSTPF150
E     D10 .......... .......... .......... .......... .......... .......... .......... .......... ..........
B=B1  D4  .......... .......... ..E....... .......... .Q........ ...VSVDVV. .......L.. .......... ..........
C     D1  .......... .......... ..E....... .......... .Q........ ...VSVDVV. .......L.. .......... ..........
A1    D2  .......... .......... ..E....... .......... .Q........ ...VSVDVV. .......L.. .......... ..........
A2    D3  .......... .......... .......... .......... .DG....... ...VSVDVV. .......L.. .......... ..........
A3    D16 .......... .......... ..E....... .......... .Q........ ...VSVDVV. .......L.. .......... ..........
A4    D21 .......... .......... ..E....... .......... .Q........ ...VSVDVV. .......L.. .......... ..........
A5    D27 .......... .......... ..E....... .......... .Q........ ...VSVDVV. .......L.. .......... ..........
      MAJ T......L.. .......... .......Q.. .....TN... .QSV.Q.... ....S.DAI. ......LL.. ....E.EP.L
      MEX T.....V... .......... ..E......V .......... AK.V...... ...SAA.I. ...D..R... ...TA.N.SL
```

(Figure 43 - Alignment of deduced protein sequence coded by *mspC3* of *L. donovani* complex strains and the outgroups *L. major* and *L. mexicana*. See page 165.)

```
Group                  *                                                                                  +                    aa
D     INF SLYWQYFTNA SLGGYSPFLD YCPFVIGYSD GSCNQDASLA AGFFSAFNVF SDAARCIDGA FRPKNRTAAN GYYAGLCANV230
E     D10 .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... 
B=B1  D4  .P.......I .......... ........G. .......... T...G..... .......... .........D .......... 
C     D1  .P.......I .......... ........GG .......... T...G..... .......... .........D .......... 
A1    D2  .P.......I .......... ........G. .......... T...G..... .......... .........D .......... 
A2    D3  .P.......I .......... ........G. .......... T...G..... .......... .........D .......... 
A3    D16 .P.......I .......... ........G. .......... T...G..... .......... .........D .......... 
A4    D21 .P.......I .......... ........G. .......... T...G..... .......... .........D .......... 
A5    D27 .P.......I .......... ........G. .......... T...G..... .......... .........D .......... 
      MAJ PR........ .......... .........A. ........S. EE..T..... .......... ....ATNGIV KS........ 
      MEX AT........ .......... .....V..RN ........TT PDLLA..... .E........ .T......D ...TA..... 

                                                    +                                                    (gpi)
D     INF RCDTATRTYS VQVRGSMDYV NCTPGLRVEL STVSSAFEEG GYITCPPYVE VCQANVKGAK DFAGDSDSSS SAGDAADRAA310
E     D10 .......... .......... .......... .......... .......... ...G..QA.. .......... .......... 
B=B1  D4  .......... ...C...... .......... .......... .......... .......... .......... .......... 
C     D1  .......... ...C...... .......... .S........ .......... .......... .......... .......... 
A1    D2  .......... .......... .......... .......... .......... .......... .......... .......... 
A2    D3  .......... .......... .......... .......... .......... .......... .......... .......... 
A3    D16 .......... .......... .......... .......... .......... .......... .......... .......... 
A4    D21 .......... .......... .......... .......... .......... .......... .......... .......... 
A5    D27 .......... .......... .......... .......... .......... .......... .......... .......... 
      MAJ Q......... ...H..N..T .......... ....KT.... .......... ...G..QA.. ..D..... .SS....K.. 
      MEX K......... .....TNG.A ........K. .S..D...KG ..V....... ...G...A.. .....T.... ..D....KE. 
      D23A                    . S......... ....N..... .......... ...G...... .......... ..D...GK.. 61
      D23B                ..... S......... ....N..... .?........ ...G...... .......... ..D...GK.- 66
```

(Figure 43 - Alignment of deduced protein sequence coded by *mspC3* of *L. donovani* complex strains and the outgroups *L. major* and *L. mexicana*. See page 165.)

```
Group                                                                                    aa
D     INF MQRWNDRMAG LATAAMVLLG MVLSLMALVV VWLLLLTCPW WCCKFGGLPT *                       360
E     D10 T..........  ..........  ..........  ..........  ..........  *
B=B1  D4  ..........  ..........  ..........  ..........  ..........  *
C     D1  ..........  ..........  ..........  ..........  ..........  *
A1    D2  ..........  ..........  ..........  ..........  ..........  *
A2    D3  ..........  ..........  ..........  ..........  ..........  *
A3    D16 ..........  ..........  ..........  ..........  ..........  *
A4    D21 ..........  ..........  ..........  ..........  ..........  *
A5    D27 ..........  ..........  ..........  ..........  ..........  *
      MAJ IE...E....  .....T....  V.........  .....VS..R  ....V.....  *
      MEX ....S....A  ....TTL...  ........L.  .R...TSS..  C..RL.....  *
      D23A..........  .....T....  ..........  .....VS...  ....L..P.A SVTAACSPET E*  123
      D23B----------  -----T....  ..........  .....VS...  ....L..P.A SVTAACSPET E*  112
```

Figure 43 - Alignment of deduced protein sequence coded by *mspC3* of *L. donovani* complex strains and the outgroups *L. major* and *L. mexicana*. Dots are amino acids identical to the *L. infantum* sequence, dashes (-) are missing amino acids, * are stop codons and ? are undetermined amino acids. Underlined are regions for which peptides were studied as T cell epitopes (Morales *et al.*, 1997; Russo *et al.*, 1993). In one box is the second metalloprotease zinc binding site (zn) and in the other is the site where the GPI anchor should be (Ramamoorthy *et al.*, 1992). Potential N-linked glycosylation sites are depicted as described by Ramamoorthy*et al.* (1992) (+) in *L. chagasi*, by Voth *et al.* (1998) (×) in *L. major* and (✳) in both. Shaded boxes are amino acids specific to *L. donovani* in relation to *L. infantum / L. chagasi* and group E (D10). Groups are the same as those in Table XXXII (insert).

Table XXIX - Corrected genetic distances between *mspC* DNA (below) and protein (top) sequences as measured using Kimura 2-parameter models. Groups are the same as those in Table XXXII (insert).

Group	Code	Strain name	D	E	B1	B	A2	A1	A3	A5	A4	C	*L. major*	*L. mexicana*
D	Inf	IPT 1	-	0.0169	0.0459	0.0459	0.0460	0.0459	0.0429	0.0459	0.0429	0.0578	0.2071	0.2446
E	D10	VL29	0.0121	-	0.0639	0.0639	0.0639	0.0639	0.0608	0.0639	0.0608	0.0700	0.1890	0.2524
B1	D	DD8	0.0283	0.0351	-	0.0000	0.0140	0.0056	0.0028	0.0056	0.0028	0.0112	0.2071	0.2524
B	D4	Patna 1	0.0273	0.0341	0.0009	-	0.0140	0.0056	0.0028	0.0056	0.0028	0.0112	0.2071	0.2524
A2	D3	MRC74	0.0312	0.0400	0.0121	0.0112	-	0.0084	0.0112	0.0140	0.0112	0.0255	0.2035	0.2602
A1	D2	MRC3	0.0292	0.0380	0.0084	0.0074	0.0056	-	0.0028	0.0056	0.0028	0.0169	0.2035	0.2524
A3	D16	Ndandu 4A	0.0283	0.0370	0.0074	0.0065	0.0084	0.0028	-	0.0028	0.0000	0.0140	0.2071	0.2485
A5	D27	Ayele 8	0.0283	0.0370	0.0074	0.0065	0.0084	0.0028	0.0019	-	0.0028	0.0169	0.2108	0.2485
A4	D21	LRC-L57	0.0273	0.0360	0.0065	0.0056	0.0074	0.0019	0.0009	0.0009	-	0.0140	0.2071	0.2485
C	D1	HU3	0.0350	0.0380	0.0084	0.0074	0.0187	0.0131	0.0121	0.0121	0.0112	-	0.2145	0.2524
	Maj	*L. major*	0.1271	0.1214	0.1372	0.1372	0.1382	0.1382	0.1418	0.1418	0.1406	0.1395	-	0.3212
	Mex	*L. mexicana*	0.1347	0.1415	0.1407	0.1418	0.1475	0.1440	0.1428	0.1428	0.1416	0.1475	0.1757	-

Legend: 0-0.0050 | 0.0051- 0.0100 | 0.0101- 0.0200 | 0.0201- 0.0300 | 0.0301- 0.1000 distance

6.5. Phylogenies of the *Leishmania donovani* complex

The *mspC3* sequences obtained from strains of the *L. donovani* complex were aligned with published homologues for two outgroup species: *L. major gp63-6* and *L. mexicana gp63-C1*, from GenBank (respectively AF039721 and X64394) (Fig. 41).

In all phylogenetic analyses of *mspC*, rooted to *L. major* and *L. mexicana* (Fig. 44), the *L. donovani* complex was monophyletic. The *L. infantum / L. chagasi* (group A) were also monophyletic; they formed a branch of indistinguishable strains. The *L. donovani* formed a paraphyletic group (did not include all taxa in the clade) because branch E was closer to or before the *L. infantum* branch. All branch lengths were positive by neighbour-joining (Fig. 44) and UPGMA, but not Fitch-Margoliash, from Kimura 2-parameter corrected genetic distances. In the other phylogenetic methods tested, topologies were similar, except by maximum likelihood which placed group E at the base of the *L. donovani* tree. Bootstrap values were high for the main branches, especially for neighbour-joining. The *L. infantum / L. chagasi* (group D) were grouped with group E strains at a long distance from the remaining *L. donovani* complex strains. However, groups E and D were not closely associated and some equally parsimonious trees had a trifurcation at the base of the *L. donovani* complex, with group D (*L. infantum / L. chagas)*, group E and the other *L. donovani* as main branches. The latter were more polymorphic than *L. infantum* and two subgroups could be defined by both visual inspection (Fig. 41) and tree analysis of the sequences (Fig. 44): one comprising LON 41 (Indian) (group B) and LON 46, 48 and 50 (group C) and the second, more polymorphic, which included LON 44, 45, 51, 56 (group A).

Protein sequences generated similar trees as DNA sequences with neighbour-joining and parsimony. Resolution of strains in groups A to C, however, was much lower than with DNA sequences, failing to isolate a clade A. All branch lengths were positive by neighbour-joining (Fig. 44) and UPGMA, but not Fitch-Margoliash.

Correlation with zymodeme type, as typed in this project (Chapter 3) was observed for all main groups, although stronger for clade D (*L. infantum / L. chagasi)* and group E strains (LON 42/52; ILM 10). One branch included closely related zymodemes ILM 5, 6, 7 (only distinguishable by ASAT), and another two ILM 3 strains. The LON 41 (ILM1) strains, which formed group B, were in one branch except for DD8 (see discussion). ILM 4 and ILM 8 strains presented much more diversity and were divided into three very closely related branches.

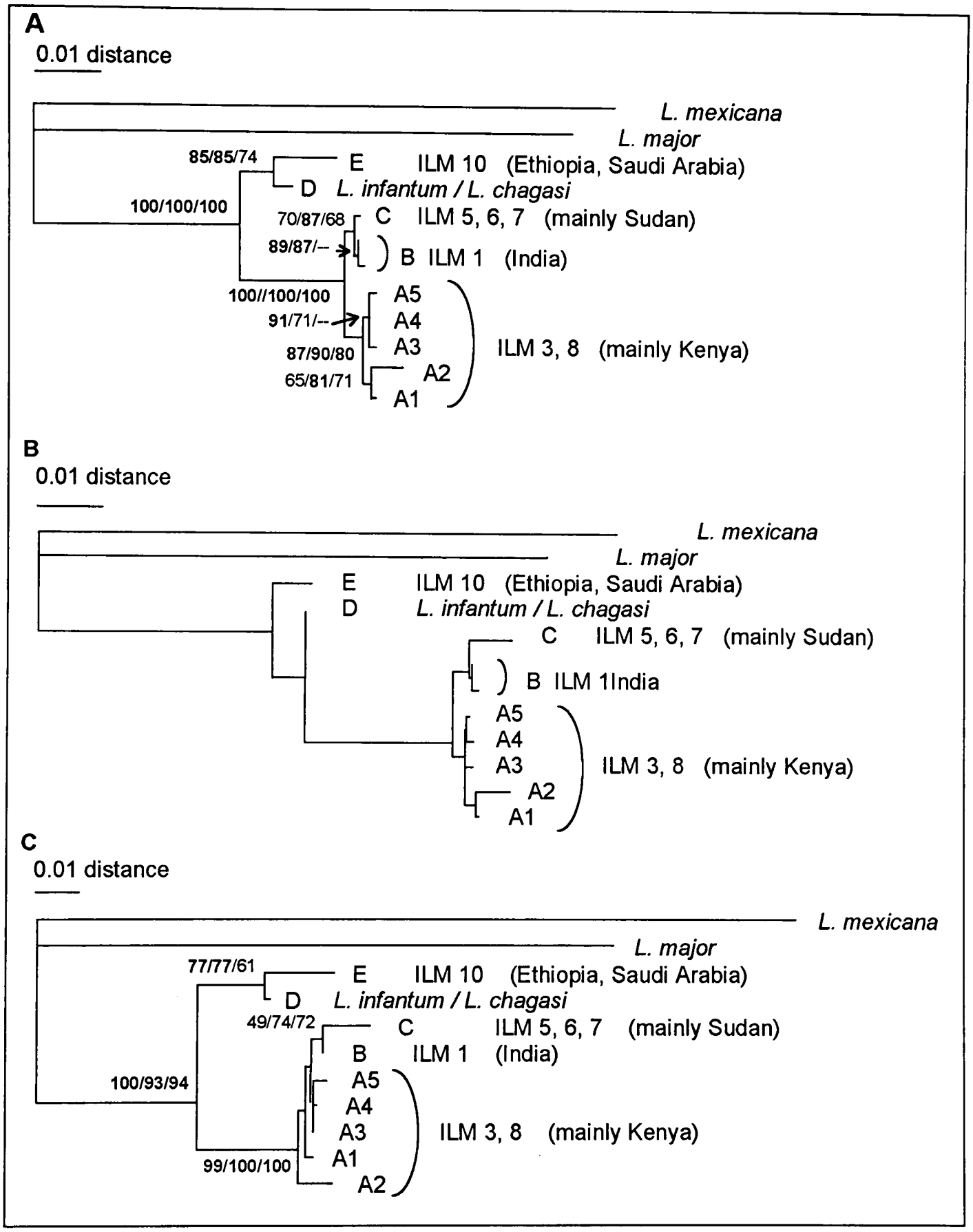


Figure 44 - Neighbour-joining dendrograms of *mspC3* A) DNA and C) protein sequences in the *L. donovani* complex, with *L. major* and *L. mexicana* as outgroups, from Kimura genetic distances (Table XXIX). *Leishmania donovani* complex strain groups were numbered according to the list previously presented (see 6.4.). Maximum likelihood, maximum parsimony and Fitch-Margoliash (SQ=0.1237, ASD= 3.251) produced similar trees, except for maximum likelihood which tree is shown in B) (Ln Likelihood = -2869). Bootstrap values from 1000 replicates in A) and from 100 replicates in C) are shown for neighbour-joining / Fitch-Margoliash / maximum parsimony.

6.6. Partial GP63 ITG/CS sequences

Partial sequences of the transcribed non-translated 3' tail of the *mspC* gene (GP63 ITG/CS) were also obtained (Fig. 46). In 147 base pairs (bp), 6 (4.1%) polymorphic sites were positively identified within the *L. donovani* complex, except for strain Sukkar 2 (D23), which was much more polymorphic.

None of the obtained sequences had the 19bp repeat described in the first published *L. chagasi mspC* (Fig. 45).

```
L. chagasi   2122 tagcggcagtatgtcctcatgtcggcagtatgtcctcatgtcggcagc
L. infantum  1183 tagcggcagtatgtcctcatgt----------------cggcagc
L. infantum  1183 tagcggcag----------------tatgtcctcatgtcggcagc
```

Figure 45 - An alignment of the published *L. chagasi mspC* sequence (M80671) and the *L. donovani* complex sequence obtained in this project (*L. infantum* IPT 1, AJ010234). Two alternative *L. donovani* complex alignments are shown, highlighting the 19bp repeat in *L. chagasi*. Identical sequences at the extremes of the repeat are double underlined. Nucleotide positions are from GenBank sequences .

```
Group
      mspC  3gttacgacggtcggcgagcttgaaacggagtggagaggatggccatcggcgagaaggccgcgacgaagtgtagcggcagtatgtcc
D     INF    gttacgacggtcggcgagcttgaaacggagtggagaggatggccatcggcgagaaggccgcgacgaagtgtagcggcagtatgtcc
E     D10    ...............................................................................
B1    D      ................g.c.....................................................c.....
B     D4     ................g.c.....................................................c.....
C     D1     ................g.c.....................................................c.....
A1    D2     ................a.....................................................c.....
A2    D3     ................a.....................................................c.....
A3    D16    ................a.....................................................c.....
A4    D21    ................a.....................................................c.....
A5    D27    ................a.....................................................c.....
      D23    .....g...c.t..tc..c.........a..t..............g...a....a.............c.....
```

(Figure 46 - Alignment of gp63 ITG/CS sequences, following the stop codon in *mspC*. See next page.)

```
Group
     mspC 80tcatgtcggcag(...)ccatcgctgcacacgtgcgccggagtgttattattattatgcttcgctgtcgctcgtcttctccggtcca
D    INF    tcatgtcggcag     ccatcgctgcacacgtgcgccggagtgttattattattatgcttcgctgtcgctcgtcttctccggtcca
E    D10    ............     ..................................................
B1   D      ............     ........t.....................................n.......n.
B    D4     ............     ........t.....................................
C    D1     ............     ........t.....................................
A1   D2     ............     ........t............
A2   D3     ............     ........t..............g..............n.nn......n.
A3   D16    ............     ........t..............g.....................................
A4   D21    ............     ........t..............g.....................................-....
A5   D27    ............     ........t..............g.....................................
     D23    c....c....g.     ......g.....................c.....................c.....................n.


mspC  191 ccttcggagtctgagctgggtatgtgtttgggtggatggggctgtttcacagcctccgtctccatccctttcctcccttctgccg
INF       ccttcggagtctgagctgggtatgtgtttgggtggatggggctgtttcacagcctccgtctccatccctttcctcccttctgccg

mspC  276 tgggtgtgtgcttgccgccgataagcaccttaagttcttggtcagccatcgcgcaatggccagatgatcatcaaggtcagcgggc
INF       tgggtgtgtgcttgccgccgataancaccttaanttcttggtcagccatctcccaanggccanatnatcatcanngtcancgggc

mspC  361 ggctctcgacgcatctcgcggtgccgtcgtgccacaatgtgtgcctctgttacgcttttctgcttcgtgctgttttgctatt
INF       ggctctcnanncatctcgcggtgccntcntgccacaangtgtgcctctgttaccctntctctgcnnectgctgtcttgcnatt
```

Figure 46 - Alignment of gp63 ITG/CS sequences, following the stop codon in *mspC*, and compared with the published *L.* chagasi 3' end sequence (mspC; S81769) (Streit *et al.*, 1996). The stop codon of strain Sukkar 2 (D23) (double underlined) is shown and a partial sequence of the coding region for comparison purposes. Dots are bases similar to the *L. chagasi* PP75 *mspC* sequenced in this project (INF) and dashes (-) were used for optimal alignment. Further gp63 ITG/CS sequence obtained for *L. chagasi* PP75 is shown below. (...) represents the 19 nt insertion present in *mspC*. Groups are the same as those in Table XXXII (insert).

6.7. Discussion

6.7.1. Comparison of *mspC* sequences with *Leishmania chagasi mspC*

The sequence of the original *L. chagasi mspC* was not identical to that found in the *L. donovani* complex, or in *L. major* or *L. mexicana*.

The 3 base gap found in the original *L. chagasi mspC* sequence (M80671):

1450 `catcccttggg--ctactcgc-atttct` 1474

may have been a sequencing error. The high number of repeated Gs may cause compression of bands and be more difficult to read. The authors may have speculated on the actual sequence to produce a working gene sequence, because a true mutation, although not impossible, would require two separate deletion events, each of which would have to be an out of frame mutation.

Similarly, in positions 19 and 20, all other sequences have GG (*L. mexicana* CG), while the original *L. chagasi* has CC.

The strain used in the original work is not one well known or even a reference strain. It may be that it has been cultured in the laboratory for a long time without selective pressure and thus may have accumulated mutations. Alterations during cloning are also possible. This case advises against the use of non-reference strains and stresses the need for including comparative sequence analysis with other strains of the same species when describing a new gene.

6.7.2. Comparative analysis of *Leishmania donovani* complex *mspC* genes

6.7.2.1. Typing markers

Upon *mspC3* sequence analysis, *L. infantum* and *L. chagasi* strains were indistinguishable, in contrast with variability between *L. donovani* strains. It is a strong argument against separation of *L. infantum* and *L. chagasi* into distinct species.

Sequence identity of *L. infantum* and *L. chagasi* strains suggested that it might be possible to identify markers present in all *L. infantum* and *L. chagasi* for specific typing of *L. infantum / L. chagasi*. Sites 1011 and 1020 are *L. donovani* specific and some other sites are *L. donovani* specific except for group E (LON42/52). Sites 634 and 932 are LON 42/52 specific but there are no single *L. infantum* specific markers because of the high similarity of group E with *L. infantum* and some similarities with *L. major* and *L. mexicana*, which suggests that, upon analysis of other related species or *L. donovani* complex strains, it may be impossible to define a single diagnostic marker within this gene for either species or group. Instead it may be necessary to rely on groups of markers for specific diagnosis / typing. These markers could be

achieved by designing probes to the desired regions or by combinations of primers to design specific nested-PCRs.

6.7.2.2. Gene and function

There was a higher percentage of amino acid (7.75%) than nucleotide polymorphisms (5.3%) throughout the analysed portion of the gene, which is a consequence of the large number of polymorphisms in first and second codon positions, which were mostly non-silent. Some regions did not have amino acid polymorphisms, whilst others are extremely divergent. Although there were 33 non-silent substitutions, in four cases there was more than one polymorphic site per codon, and thus only 28 amino acid polymorphisms.

Although the third codon position was more polymorphic than the first or second, the difference between the observed and the expected distributions if distribution of mutation was equal among sites, was not statistically different to a degree of 10% by the χ^2 test (Table XXX), ie., the difference in polymorphism frequency between codon positions 1, 2 and 3 is not statistically significant. No substitution restraints were identified regarding codon position or synonymy of base change, given that first and second base substitutions were not silent, and thus, regions with genetic polymorphisms in *mspC* did not seem to be targets for conservative selection.

Table XXX - χ^2 statistics of distribution of base changes over the three codon positions.

Position	1	2	3	total
Observed	18	15	23	57
Expected	18.3	18.3	18.3	56
$H_0 : n_1 = n_2 = n_3$	$H_1 : n_1 \neq n_2 \neq n_3$			
$\chi^2 = \Sigma [(O - E)^2 / E]$	2 degrees freedom		$\chi^2 = 1.33 < p=0.5$	

It is tempting to speculate over the large percentage of nucleotide polymorphisms producing amino acid substitutions. *MspC* does not seem to be a pseudogene, because findings of only in frame, and no out of frame, mutations argue against this hypothesis. Similarly, distribution of polymorphisms does not seem entirely random and there are many conserved motifs among *L. donovani*. It may be that *mspC* is mutation tolerant or even that mutations may often be beneficial rather than harmful. Because GP63 is a surface protein, the parasite may benefit from diversity in some antigenic regions.

The most promising T cell epitope for yielding protection against leishmaniasis (peptide 7- shown in bold in Fig. 47) produced similar effects whether given in the *L. major* or in *the L. chagasi* form, despite major sequence differences, and divergent homologous sequences seem to encode for immunogenic peptides (Fig. 47). The first region of the Zn binding site may also have important antigenic properties (Yang *et al.*, 1993), but it is located upstream of the region sequenced in this work. Published data on antigenic epitopes are rather difficult to interpret, since different peptides of different lengths have been used in different experiments. It may be that position of certain sequences may influence antigen presentation and hence the immune response. Except for the case with peptide 7, it is not known whether sequence differences between strains are important or not for the immune response.

Three polymorphic regions were identified within the seemingly immunogenic region shown in Fig. 47. Although different genes may have been compared in the several antigenic epitopes studied, there are some common species specific amino acid patterns and it is tempting to relate amino acid motifs to the immune response elicited by each parasite. Especially in the third box, the *L. donovani* complex has NRTAA$^{N}/_{D}$GY and *L. major* genes have AT$^{D}/_{N}$GIVKS. Is it possible that one of the motifs elicits a Th1 response and the other a Th2 response, thus influencing the outcome of infection and participating in the mechanisms by which one species causes the milder CL form and the other the dangerous VL. *Leishmania mexicana*, however, has a sequence similar to VL species. *Leishmania* pathogenesis is far from simple, however, and this subject deserves further investigation.

Differences in gene sequence may have implications in antigenic and virulence properties. Contradictory results on the study of GP63 may reflect differential regulation or amino acid composition of the protein. Although not exclusive of this protein, the comparative study of GP63 sequences and properties might provide some light on the striking differences found among *Leishmania* species regarding immunogenicity and clinical presentation. I believe that this analysis shows how important it is to integrate functional studies with comparative genetics.

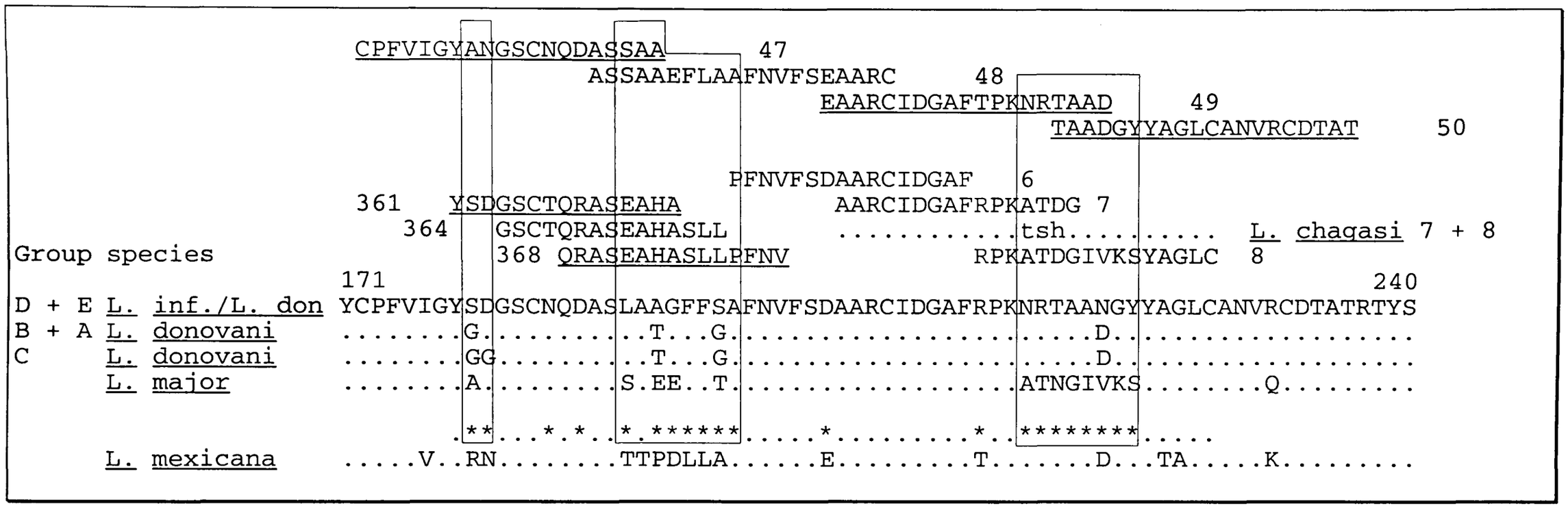


Figure 47 - Alignment of the most studied region for GP63 antigens. Where *L. major* sequences were studied a sequence comparison is shown. * represents polymorphic sites. The most variable regions are highlighted by boxes. Epitopes 47 to 50 were studied by Morales *et al.* (1997) and were from *L. infantum*, 361 to 364 by Soares *et al.* (1994) and 6 to 7 by Russo *et al.* (1993) from *L. major*. Epitopes 7 and 8 were shown to stimulate T cell proliferation in leishmaniasis patients. Epitopes 47, 49 and 50, underlined, produced high FAST-ELISA titres. Epitope 364 was shown to produce a protective Th1 response, while 361 and 368 (underlined) were shown to produce an exacerbating Th2 response.

6.7.3. Phylogenies of the *L. donovani* complex from *mspC3*

MspC3 was polymorphic within *L. donovani*, but not *L. infantum / L. chagasi*. High bootstrap values strongly support the main branches. As far as we are aware, this is the first time that the sequence of a GP63 gene has been shown to be of value for inference of phylogenies in *Leishmania*.

As shown in the dendrogram, both *L. infantum* and the *L. donovani* LON 42/52 groups were closer to the base of the tree, than most of the *L. donovani*. Position of the LON 42/52 was uncertain in this phylogeny, but they seemed at least as divergent as *L. infantum* (*L. chagasi*) is from other *L. donovani* strains.

L. donovani complex strains could be grouped according to isoenzyme types (Fig. 44), which corroborates the initial hypothesis that phylogeny of GP63 may correlate with biological properties of the parasites. *MspC* is thus a strong candidate for the development of molecular typing methods to supplement isoenzyme typing, although less variable. Possible typing methods could be nested-PCR or molecular beacons.

Strains DD8 (D) and MRC74 (D3), although similar to strains of their respective zymodeme groups, had singular sequences. DD8 also has two different sized mini-exon alleles (Chapter 5). It is possible that most strains are essentially clones circulating in the same form throughout epidemics. Others, like DD8 and MRC74 may have been under different selective pressures or may be strains of clonal lines less represented in the populational gene pool. It is known that PCR amplification may introduce errors in sequences, but it is unlikely that they should be introduced preferentially in one strain and not in others.

The phylogenetic analysis using the protein sequence rather than the DNA sequence produced similar results, except for the clade C-A. Despite this, association of type C (strains ILM 5/6/7) and Indian *L. donovani* (type B) was still strong. This confirms that protein sequences are good tools to infer phylogenies, except from closely related sequences. Despite this constraint, it is useful to verify that purely genetic distances correlated with phenotypic distances. Again, this effect may be due to the observed high rate of non-synonymous polymorphisms and suggests a strong positive selection (favouring diversity) acting upon this gene or in some regions, instead of negative selection (favouring conservation).

MspC genes are sufficiently conserved to assume that the trees obtained reflect the true phylogeny of the gene, which also reflects phylogeny of the parasites, and it is, thus, possible to speculate on the history of the *L. donovani* complex. In the *mspC* phylogenies the *L. donovani* complex was only slightly more closely related to *L. major* than to *L. mexicana*, which was more evident at the protein level. Although

analysis of more strains and input of a more distant outgroup would be necessary, data here were compatible with the hypothesis of radiation of *L. mexicana*-like parasites into the Old World with quick radiation of species in the new environment.

At least two main lines of *L. donovani* complex evolution emerged: one with *L. infantum* (group D) and ILM 10 strains (group E) and the other with clade A-C. Both *L.* infantum and group E were closer to the outgroups but it was not clear whether those formed a clade or were two ancient groups. Indeed, *L. infantum* and group E were more distant from each other than any sequences within clade A-C which was much more polymorphic. Phylogeny within clade A-C was not well resolved and it may be that addition of other strains may complicate it. *Leishmania infantum* and group E seemed to have undergone little evolutionary change as far as the *mspC* gene is concerned whilst clade A-C has undergone fast radiative evolution. Whether the evolutionary patterns are associated with vectors, with reservoirs or with simple population events, such as population bottlenecks, deserves further investigation.

6.7.4. Sequence of gp63 ITG/CS

Partial sequences of the transcribed non-translated 3' tail of the *mspC* gene were also obtained. Although lower, the number of polymorphic sites (4%) found in the non-coding region were close to the 5.2% in the coding region sequenced in this work. The short sequence data did not allow for conclusions. Some parts of intergenic regions are expected to be conserved because of their function in regulation of transcription and mRNA synthesis.

Analysis of the sequences obtained indicated that the first *L. chagasi mspC* sequence had a 19bp repeat in the non-coding region. Figure 48 shows how this duplication may have occurred either in the parasite or during the cloning-sequencing process. The six base repeat present twice in the initial sequence may be due to slipped annealing, which would initiate a second amplification of the sequence that is repeated. The duplication may have become stable through a DNA repair mechanism involving new DNA synthesis from the single strand DNA in the loop or after chromosome replication in the next mitosis event.

Mispaired crossing-over might produce a similar effect; it should have been found, however, a shorter allele with a deletion of the sequence repeated.

6.7.5. Strain Sukkar 2

Strain Sukkar 2 (D23) was different from any strain studied here and possessed a few remarkable features. Two PCR products were identified, the shorter fragment of which revealed a 48bp in frame deletion. The deletion may have

occurred by formation of a long loop with pairing of 7bp during transcription (Fig. 49A). Six of these pairings are C-G, which are very strong, thus stabilizing the loop enough to allow transcription. A new DNA molecule with the deletion would be produced, either by a second replication of the shorter strand in the next mitosis event or by DNA repair mechanisms (see 6.7.4.). Alternatively, the deletion might have occured through slippage involving the 6bp repeat sequence in the template strand which could be stabilized by a short loop with four base pairings, three of which C-G (Fig. 49B). Deletions may also occur through mispaired crossing-over, but it would most likely not produce the truncated repeat 'gcg' present in the deletion, and would cause a duplication in the sister chromosome. This duplication would have to be observed if *Leishmania* reproduction is assexual and only mitotic.

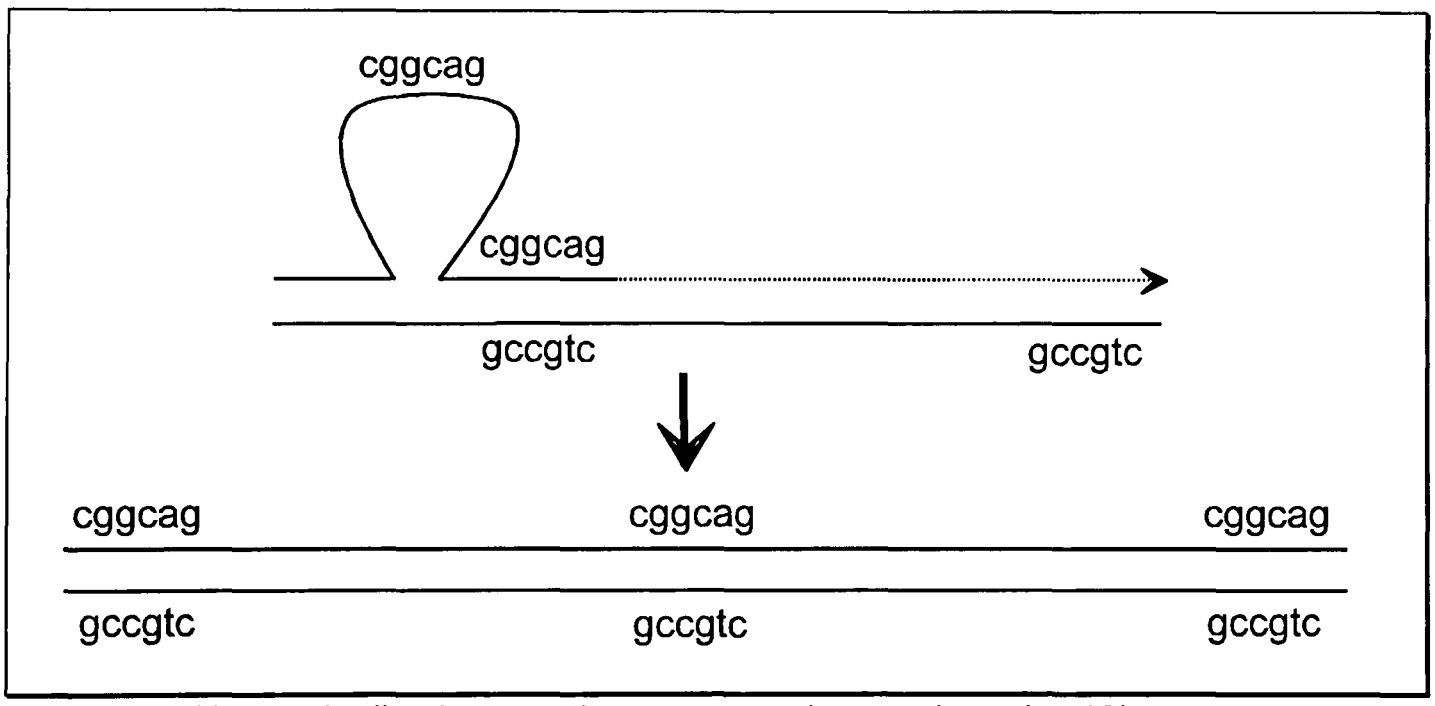


Figure 48 - How a duplication may have occurred to produce the 19bp repeat present in the published *L. chagasi mspC* sequence. The lower strand is the sequence present in all *L. donovani* strains studied in this project and the upper strand is the newly synthesized strand after slippage of the new DNA molecule during synthesis. Below is the resulting DNA after a second round of amplification or DNA repair with synthesis of DNA complementary to the single strand DNA of the loop.

Remarkably, both SUKKAR 2 sequences had a polymorphism in what is a stop codon for every other studied strain, which coded for a serine residue. The mutation may not affect the function of the gene, since there is a second stop codon in both fragments, only 11 amino acids downstream. This second stop codon is not present in any of the studied *L. donovani* complex strains and it is located in a region conserved within the *L. donovani* complex. There are no other stop codons downstream for a large distance.

It is intriguing how such polymorphisms, distinctive from any other strain (or species) studied in this work, become present in both fragments. This may indicate that these events precluded the deletion event. Although gene duplication may have occurred, no other fragments were identified. Karyotypic analysis might shed some light on whether or not the two alleles were in the same locus, but itwas beyond the scope of this work.

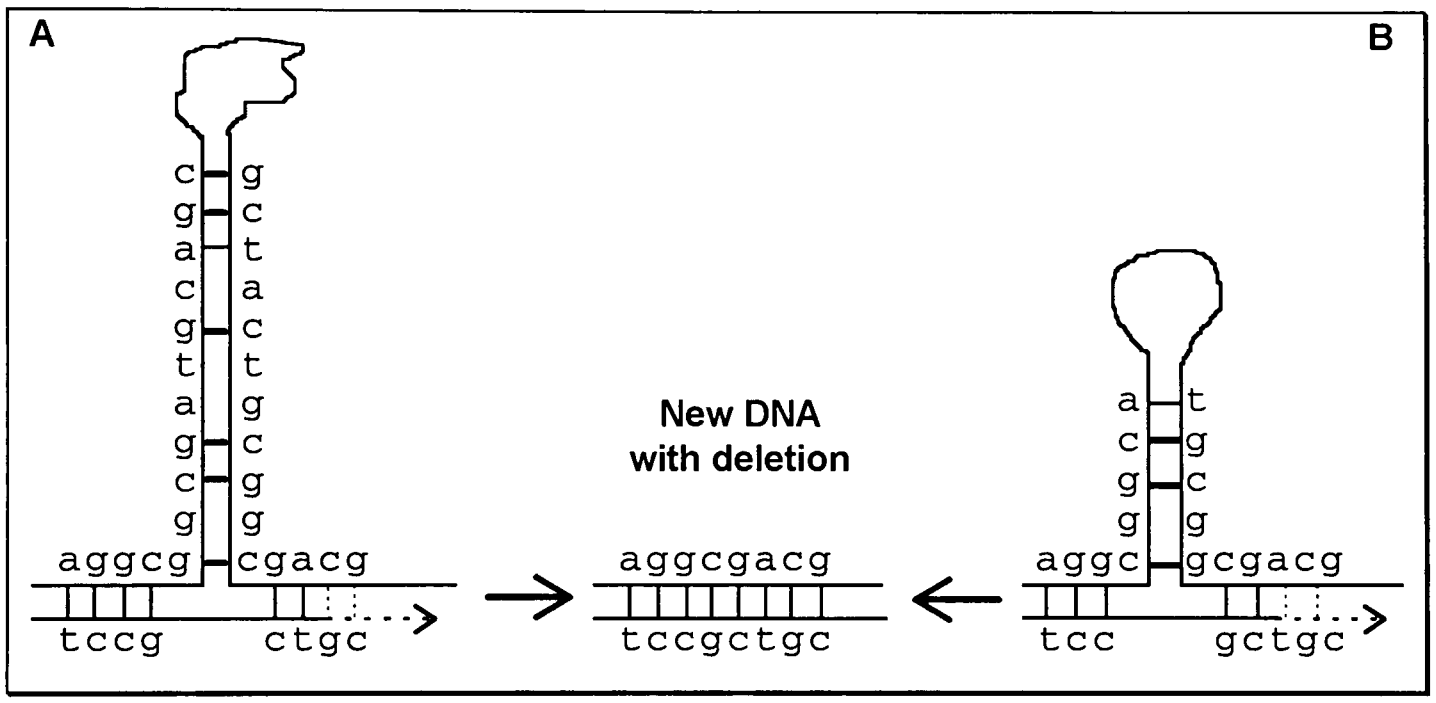


Figure 49 - The two best explanations of how a deletion may have occurred in a newly synthesized DNA strand in strain Sukkar 2 (D23). A) A long loop is formed, thus causing replication to continue through the base of the loop. B) A short loop stabilizes a slippage of the template strand during replication. Both these alternatives explain why sequence 'gcg' is present only once in the shorter fragment but a double repeat 'gcggcg' is be present in both flanks of the deletion region in the longer fragment. The dotted arrow shows the direction of synthesis of the new DNA strand and the thick small lines show C-G base pairing. A new DNA molecule with a deletion will eventually be produced (see text).

7. Overall Phylogenetic Approach to the *Leishmania donovani* complex

7.1. Potential phylogenies of the *Leishmania donovani* complex

7.1.1. Phylogenies from pooled RFLP and RAPD data

RAPD and RFLP data were pooled, since both data sets produced trees with similar genetic groups. By pooling RFLP and RAPD data, phylogenies were constructed for the *L. donovani* complex using clustering or maximum parsimony methods. All distance and parsimony based trees discriminated between five main clades or groups of strains with a possible sixth (Fig. 50), which had strong correlations with geographical origin or zymodeme (Fig. 52):

A - mainly zymodeme ILM 3; mainly Kenyan strains;

B - ILM 1; Indian strains;

C - ILM 5, 6, 7; mainly Sudan but includes isolates from several other locations, such as Ethiopia, Lebanon, Iran, Italy, Portugal;

D - *L. infantum* (ILM 9 and related zymodemes); several locations in Mediterranean countries to China, and Latin America (not all strains were shown here);

E - ILM 10; Ethiopian and Saudi Arabian strains;

F - ILM 16; China. Here represented by one strain only, because, although usually associated with, it was not sufficiently close to group E (ILM 10).

All groups were robust upon bootstrap analysis of Wagner parsimony (Fig. 50), which produced the shortest cladograms (642 total steps), followed by Dollo (772 total steps) and polymorphism (1351 total polymorphisms). Groups and their relationships were congruent in all phenograms [neighbour joining, UPGMA (CC= 0.94) and Fitch-Margoliash (SQ= 1.17; ASD= 3.44)] and cladograms (Wagner, Dollo or polymorphism parsimony). The only exception was that in the Fitch-Margoliash tree group A (Kenyan strains) was closer to group C, than to group B (Indian strains), however the distances in a Fitch-Margoliash tree between groups A, B and C were negative and bootstrap values were lower than 70% for either combination AB, BC or AC. These three groups thus emerge as a tricotomy.

The five groups of strains could also be defined at 65% of the maximum distance within the *L. donovani* complex (Table XXXI), as with single RFLP analyses. Increasing the discrimination level to 75% improved definition of groups C and D without compromising overall definition, except for strains D2 and D3.

Several degrees of geographical diversity were apparent among the groups. The most geographically restricted was group B (Indian strains) followed by groups A

and E (in Kenya/Ethiopia and in Ethiopia/Saudi Arabia, respectively). The most disperse was group D (*L. infantum*), in both OW and NW, followed by group C, in Mediterranean / North African countries. Although group F was here represented by one strain only, it is probable that it is restricted to China.

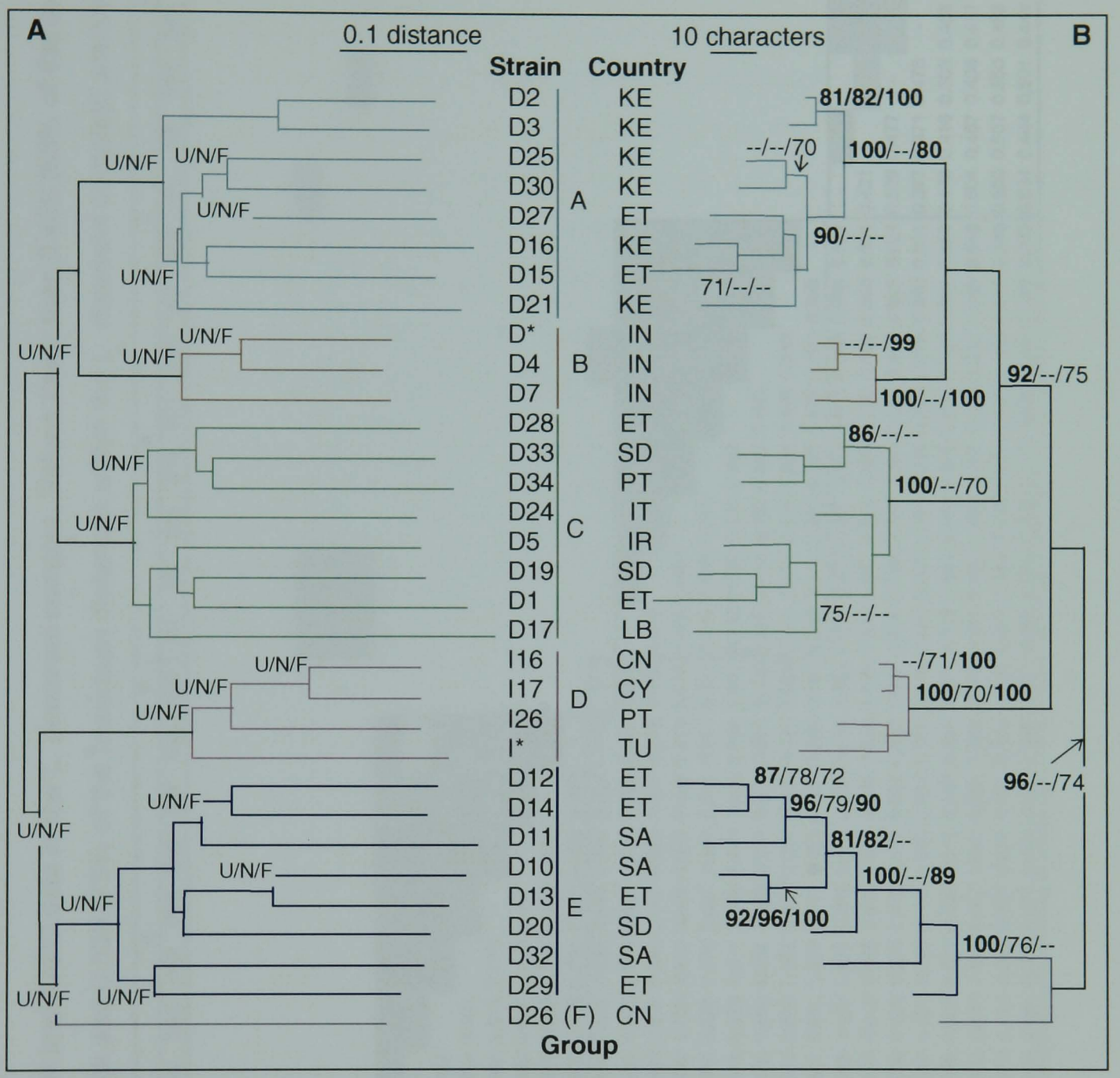


Figure 50 - Phylogenetic analysis of the *L. donovani* complex based on pooled RFLP and RAPD data. A) Neighbour-joining tree compared with B) Wagner consensus tree (642 steps). Branches present in all UPGMA / neighbour-joining / Fitch-Margoliash (U/N/F) are shown in A) and bootstrap support higher than 70% (higher than 80% in bold) for Wagner / Dollo/ polymorphism (W/D/P) parsimony are shown at branches in B). Countries are: CN - China; CY - Cyprus; ET - Ethiopia; IN - India; IR - Iran; IT - Italy; KE - Kenya; LB - Lebanon; PT - Portugal; SD - Sudan; TU - Tunisia. Groups are the same as those in Table XXXII (insert): A - ILM 3; B - ILM 1; C - ILM 5/6/7; D - *L. infantum*; E - ILM 10; (F) - ILM 16. *WHO reference strain.

Table XXXI - Jaccard distances from pooled RFLP and RAPD data in the *L. donovani* complex. Values lower than 0.459 (65% of the maximum distances within the *L. donovani* complex) are shaded black and 0.530 (75% of the maximum distances within the *L. donovani* complex) are shaded grey, on top.

Group		D				B			E								F	A								C							
	Strain	I	I16	I17	I26	D	D4	D7	D10	D11	D12	D13	D14	D20	D29	D32	D26	D2	D3	D15	D16	D21	D25	D27	D30	D1	D5	D17	D19	D24	D28	D33	D34
D	I	-																															
	I16	0.372	-																														
	I17	0.365	0.174	-																													
	I26	0.387	0.292	0.302	-																												
B	D	0.608	0.599	0.603	0.608	-																											
	D4	0.606	0.589	0.593	0.606	0.230	-																										
	D7	0.638	0.610	0.614	0.620	0.320	0.329	-																									
E	D10	0.658	0.634	0.649	0.613	0.655	0.653	0.648	-																								
	D11	0.677	0.671	0.674	0.645	0.675	0.662	0.663	0.419	-																							
	D12	0.652	0.656	0.654	0.623	0.606	0.610	0.604	0.458	0.427	-																						
	D13	0.640	0.632	0.630	0.605	0.643	0.640	0.641	0.289	0.454	0.431	-																					
	D14	0.626	0.617	0.627	0.595	0.616	0.607	0.608	0.414	0.364	0.314	0.429	-																				
	D20	0.622	0.594	0.598	0.591	0.630	0.628	0.646	0.408	0.430	0.459	0.400	0.402	-																			
	D29	0.600	0.569	0.566	0.587	0.588	0.572	0.594	0.483	0.501	0.500	0.448	0.506	0.401	-																		
	D32	0.632	0.618	0.616	0.621	0.617	0.614	0.621	0.512	0.537	0.529	0.490	0.526	0.439	0.400	-																	
F	D26	0.610	0.573	0.570	0.591	0.579	0.569	0.598	0.604	0.608	0.602	0.596	0.586	0.530	0.482	0.544	-																	
A	D2	0.606	0.574	0.571	0.586	0.499	0.475	0.515	0.663	0.662	0.629	0.646	0.620	0.634	0.586	0.632	0.561	-																
	D3	0.614	0.576	0.573	0.594	0.500	0.486	0.517	0.660	0.669	0.624	0.642	0.615	0.623	0.607	0.634	0.563	0.243	-															
	D15	0.628	0.606	0.617	0.616	0.563	0.544	0.554	0.655	0.648	0.638	0.670	0.610	0.637	0.645	0.673	0.613	0.473	0.484	-														
	D16	0.617	0.587	0.591	0.591	0.557	0.546	0.532	0.681	0.685	0.655	0.675	0.647	0.660	0.622	0.647	0.575	0.434	0.446	0.383														
	D21	0.634	0.599	0.596	0.609	0.568	0.565	0.559	0.696	0.706	0.638	0.671	0.652	0.654	0.633	0.658	0.600	0.446	0.436	0.475	0.400	-												
	D25	0.615	0.579	0.576	0.589	0.540	0.529	0.539	0.697	0.696	0.663	0.692	0.660	0.667	0.620	0.661	0.529	0.388	0.414	0.433	0.336	0.403	-											
	D27	0.625	0.616	0.614	0.613	0.567	0.549	0.551	0.678	0.672	0.657	0.662	0.637	0.667	0.630	0.649	0.563	0.404	0.417	0.436	0.394	0.430	0.343	-										
	D30	0.610	0.580	0.577	0.591	0.564	0.546	0.555	0.690	0.690	0.661	0.675	0.647	0.660	0.604	0.636	0.553	0.398	0.412	0.463	0.375	0.425	0.304	0.368	-									
C	D1	0.622	0.608	0.605	0.579	0.504	0.509	0.527	0.639	0.658	0.608	0.620	0.605	0.637	0.604	0.630	0.602	0.556	0.579	0.607	0.589	0.613	0.568	0.585	0.569	-								
	D5	0.638	0.630	0.628	0.603	0.543	0.547	0.526	0.658	0.672	0.619	0.636	0.628	0.657	0.648	0.650	0.619	0.589	0.597	0.624	0.607	0.630	0.580	0.590	0.601	0.431	-							
	D17	0.696	0.681	0.679	0.655	0.629	0.620	0.602	0.669	0.658	0.630	0.674	0.633	0.673	0.690	0.687	0.648	0.638	0.634	0.598	0.599	0.617	0.610	0.602	0.624	0.529	0.477	-						
	D19	0.630	0.628	0.626	0.606	0.567	0.564	0.551	0.618	0.650	0.597	0.592	0.595	0.633	0.624	0.626	0.616	0.606	0.607	0.627	0.616	0.633	0.614	0.600	0.591	0.397	0.371	0.478	-					
	D24	0.638	0.618	0.615	0.620	0.538	0.535	0.528	0.693	0.688	0.642	0.674	0.650	0.641	0.608	0.634	0.557	0.573	0.589	0.629	0.586	0.604	0.549	0.582	0.557	0.478	0.416	0.505	0.468	-				
	D28	0.637	0.603	0.601	0.625	0.560	0.557	0.551	0.678	0.682	0.640	0.662	0.632	0.640	0.593	0.596	0.591	0.593	0.601	0.596	0.584	0.589	0.562	0.566	0.548	0.504	0.467	0.486	0.457	0.401	-			
	D33	0.623	0.607	0.605	0.617	0.525	0.538	0.548	0.666	0.675	0.614	0.649	0.611	0.626	0.591	0.552	0.595	0.583	0.584	0.600	0.582	0.600	0.580	0.577	0.545	0.500	0.507	0.550	0.463	0.482	0.354	-		
	D34	0.630	0.602	0.600	0.618	0.550	0.563	0.557	0.677	0.701	0.651	0.667	0.654	0.650	0.641	0.588	0.603	0.605	0.593	0.608	0.576	0.608	0.581	0.592	0.583	0.534	0.444	0.501	0.444	0.453	0.382	0.362	-	

7.1.2. Comparison of DNA sequence and fragment based phylogenies.

The same main groups as in RAPD/RFLP trees could be identified in the *mspC* sequence trees, but both distances and topology of groups A (ILM 3), B (ILM 1) and C (ILM 5/6/7) were different. Position of groups D and E was also variable in relation to each other. Highly diversified groups, such as C (ILM 5/6/7) and E (ILM 10), had identical *mspC* sequences, whilst group A was the most polymorphic cluster by *mspC*.

Rooting of most *mspC* trees differed from the RAPD/RFLP UPGMA tree in that it separated *L. donovani* in groups ABC and DE, but the rooting methods were different. Whilst the RAPD/RFLP tree was rooted by UPGMA to the most divergent branch, the *mspC* tree was rooted by an outgroup. The maximum likelihood tree for *mspC* (Chapter 6, Fig. 44), however, did place group E closer to the root. Thus, it is possible that group E and group D belong to separate clades, with group E being the first to diverge and closely followed by group D.

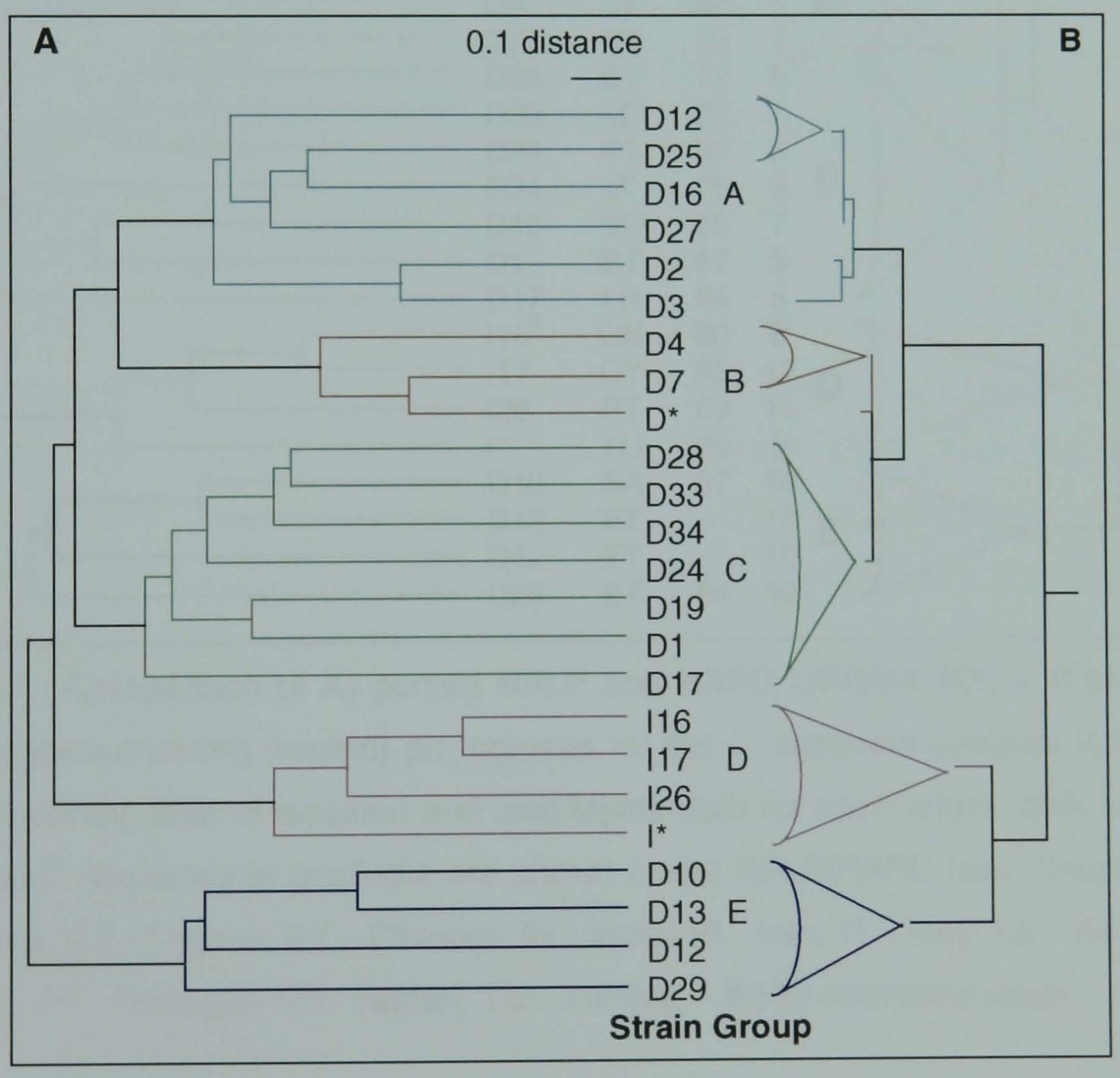


Figure 51 - Comparison of A) pooled RFLP and RAPD UPGMA (CC = 0.94) and B) *mspC* neighbour-joining (rooted) phylogenies for the *L. donovani* complex (Chapter 6). Only strains with available *mspC* sequences were used. * WHO reference strain.

7.1.3. Comparison of DNA with isoenzyme typing.

DNA based trees did not agree well with isoenzyme trees and are not shown for that reason. On one hand, ILM 8, grouped with group B within *L. donovani* in the RFLP/RAPD tree and with the *L. infantum* strains in the IEA tree. However, one strain of *L. infantum* had also been typed as ILM 8 (I1; L82). On the other hand, in the genetic trees, groups E and D are related, whilst in the phenotypic (isoenzyme) trees those groups would often be in opposite sides.

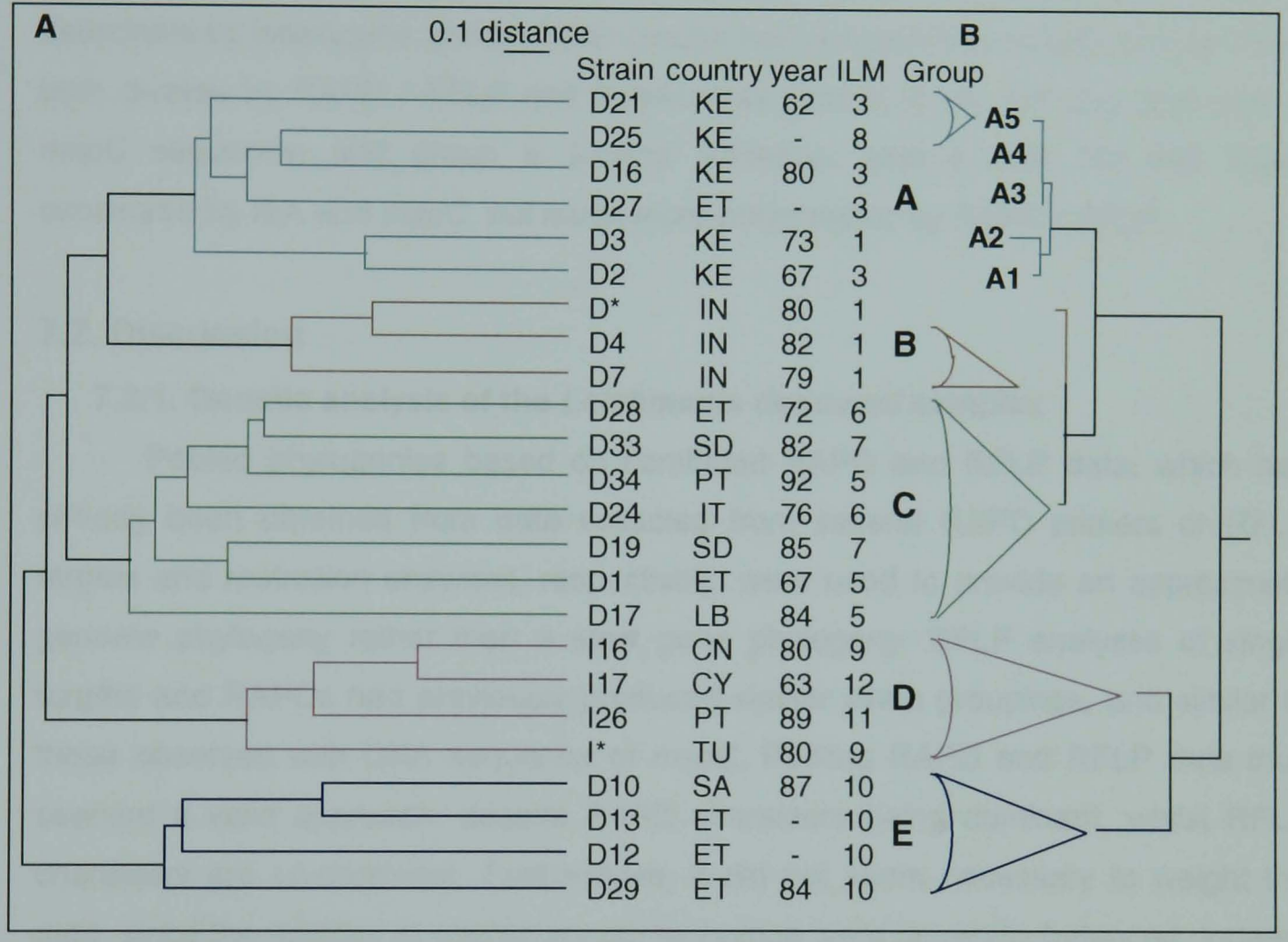


Figure 52 - Comparison of A) pooled RFLP and RAPD UPGMA (CC = 0.94) and B) *mspC* neighbour-joining (rooted) phylogenies for the *L. donovani* complex (Chapter 6) showing country, year of isolation and isoenzyme data for each strain. Only strains for which *mspC* sequence is available are shown in the RFLP/RAPD tree. Countries are: CN - China; CY - Cyprus; ET - Ethiopia; IN - India; IR - Iran; IT - Italy; KE - Kenya; LB - Lebanon; PT - Portugal; SD - Sudan; TU - Tunisia. * WHO reference strain.

Unprocessed isoenzyme data correlated well with gene (*mspC*) and genomic trees (pooled RAPD/RFLP) (Fig. 52). The exceptions were strain D3 (MRC74) in group A (ILM 3), which was typed as zymodeme ILM 1, and ILM 8 of which there were two strains in different groups. One ILM 8 strain is not shown in this tree but had already been assigned to group D (*L. infantum*) and the other is seen here with group

Chapter 3) and ILM 8 strains might be differentiated if more enzymes were used. The similarities of zymodeme typing for these strains may be due to the low number of loci studied and may have been caused by analogous, rather than homologous, alleles, which are very difficult to distinguish by IEA alone.

Congruence of methods differed according to the group. Group B (Indian strains, ILM1) was the most conserved across all methods, whilst only group A (Kenyan strains, mostly ILM 3) was diverse across all methods, however, differentiation of a sub-clade with MRC(L)3 and MRC74 within group A was not detectable by isoenzyme typing. Other groups had contradictory results. Group C was both diverse by RAPD / RFLP and isoenzymes (ILM 5, 6, 7), but only one type of *mspC* sequence, and group E (mainly Ethiopian strains, ILM 10) was highly conserved by IEA and *mspC*, but much more polymorphic by RAPD / RFLP.

7.2. Discussion

7.2.1. Genetic analysis of the *Leishmania donovani* complex

Pooled phylogenies based on combined RAPD and RFLP data, which had already been obtained from data collected from several RAPD primers or RFLP targets and restriction enzymes, respectively, were used to provide an approximate genome phylogeny rather than a strict gene phylogeny. RFLP analyses of single targets and RAPDs had previously produced similar strain groupings, and similar to those observed with DNA sequence of *mspC*. Pooling RAPD and RFLP data thus seemed a valid approach, despite RAPD characters being dominant, whilst RFLP characters are co-dominant. Furthermore, it did not seem necessary to weight the data, since the number of characters per technique were generally balanced. Indeed, the results from the pooled data were robust and compared well with those obtained with individual and other analyses, such as *mspC* sequence and isoenzyme analysis.

Groups

It was interesting to observe that not only formation of strain clusters but also relationships between groups were congruent between cluster and parsimony methods, for bootstrap values higher than 70%, when RAPD and RFLP data were pooled. Most reassuringly, in the shortest cladogram, a Wagner parsimony tree, the main branches had more than 90% support, except topology of the ABC clade. Most of the branches were also well supported in the polymorphism parsimony tree, but not so much by Dollo parsimony, perhaps because the RAPD data was less robust. Furthermore, the relative positions of the groups were stable in either pooled RFLP or RAPD/RFLP.

The groups of strains identified in the *L. donovani* complex were congruent across different methods: pooled RAPD/RFLP, *mspC* sequence and isoenzyme typing. Some zymodemes, however, were found in more than one group. ILM 8 was found in *L. infantum* (strain L82) and group A (mainly ILM 3) (strain D25, D-2), whilst ILM 1, which is characteristic the Indian group, was also found in group A (strain D3, MRC74). An interesting possibility would be that the *L. infantum* ILM 8 strain was a remnant from the ancestors of the group, and that the group A ILM 1 strain was a ancestor of Indian *L. donovani,* if *L. infantum* and *L. donovani* had both originated from group A. However, *L. infantum* was only distantly related to group A in all analyses and Indian *L. donovani* could cluster with either group A or C. It is likely, however, that the strains with zymodemes ILM 1 and ILM 8 in different groups were a case of convergent evolution, because of the low resolution power of isoenzyme analysis, in general and in this project.

Root of the *L. donovani* complex

Unfortunately, it was not possible to root the final pooled RAPD / RFLP trees, because no outgroup amplification of ITG/CS or ITG/L was possible. The structure of the *L. donovani* complex in unrooted trees was often star like, which might indicate an initial radiation or lack of a resolvable phylogeny with the gene. If this is the case, establishing the phylogeny of genetic groups of the *L. donovani* complex may be virtually impossible. The root of the complex, however could be deduced from UPGMA analyses, because it assumes a molecular clock, and from RAPD, ITS and mini-exon phylogenies, which were outgrouped.

From an overall analysis of the trees, no group is particularly close to the root of the *L. donovani* complex. The identified groups thus seem to be survivors of an unknown ancestral population, which may have been lost or highly modified. Group E appeared more often as the first clade to differentiate (with or without *L. infantum* and Wangjie 1) suggesting that group E (ILM 10) may have been the first to diverge. The *L. infantum* clade is the alternative candidate to be the first lineage to emerge from the *L. donovani* complex but ILM 10 strains (group E) were often the closest to the root, when determined. It is not clear whether *L. infantum* arose as an early branch from clade ILM 10 or as an independent line, although the latter seems to be favoured by the data. Groups A, B and C seem to be part of another clade, which may have arisen at approximately the same time as clades *L. infantum* and group E.

Although the above phylogeny may have been a result of a long branch attraction phenomenon (Felsenstein, 1988) due to a higher mutation rate in group E, it is not very likely, given the similarity of *mspC* sequences within the group and also to the closest group, *L. infantum*. The other *L. donovani* groups seemed to be much

more phenotypically diverse, as detected by isoenzyme electrophoresis, DNA sequence analysis of *mspC* but not putatively neutral diversity of intergenic regions, as detected by restriction analysis.

An alternative hypothesis for phylogenetic indeterminations is the occurrence of recombination in the origin of some groups. The analysis done here, however did not provide adequate data to resolve this question.

Phylogenies of groups

A cladogram of the genus *Leishmania* obtained from isoenzyme data (13 enzymes) by Rioux *et al.* (1990), indicated monophyly of the complex *L. donovani*, which was divided in the monophyletic sister branches *L. donovani* and *L. infantum*. ASAT was used as the determinant character for separation of the two branches. As seen above, however, *L. infantum* repeatedly appeared as a secondary clade within the *L. donovani* complex, even in parsimony trees, or, at best, one of three main clades in the root of the complex.

Groups A, B and C formed a robust clade, however the phylogenetic relationships between each group were not well determined. In particular, affinities of Indian *L. donovani* were not clear. They had lower genetic diversity than most *L. donovani* groups, including *L. infantum*, and usually formed a subgroup of either group A (ILM 3) or C (ILM 5/6/7). Strain MRC 74 from group A had a zymodeme indistinguishable here from ILM 1 (the only Indian zymodeme), although it had been differentiated by Le Blancq (1986) through MDH. Either the three groups emerged in a three clade radiation from an ancestral clade ABC population, or one of the groups originated from hybridization of strains from the other two groups. Although the present project does not make it possible to answer that question, the most likely candidate for a recombination product would be the Indian *L. donovani* group, which has characters from both the other two groups.

In summary, the most probable scenario for the phylogeny of the *L. donovani* complex is a three clade radiation near the origin of the complex, producing clades *L. infantum*, ILM 10 and ABC. The latter would have had a three clade radiation of groups A (ILM 3), B (Indian) and C (ILM 5/6/7).

Consequences of phylogeny

Species denominations.

The data from this project, thus suggest that *L.donovani*, as currently defined (Rioux *et al.*, 1990) is a paraphyletic taxa, because it does not include all organisms in the clade, since *L. infantum* seem to be part of the '*L. donovani*' clade whenever trees were rooted. Although there may be ecological and clinical reasons to maintain

L. infantum as a species, the molecular data in this work did not support species status, according to a cladistic definition of species, unless the clade with ILM 10 was also granted species status.

The fourth species of the complex, *L. archibaldi*, was not supported. Strain Gebre 1 (and all MON 82), the type strain of *L. archibaldi*, is characterised by an ASAT profile intermediate between that of *L. infantum* (MON 30) and *L. donovani* (MON 18) as shown in the IEA cladogram by Rioux *et al.* (1990), which placed MON 82 between the ancestral strains of *L. infantum* and *L. donovani*. Here, and on the analyses by Le Blancq (1986), the ASAT profiles for these strains were similar to those described by Rioux *et al.* (1990), but it was in addition obvious that '*L. archibaldi*' profiles resemble more a heterozygous profile than a true 'intermediate' or third type profile. Accordingly, Gebre 1 (ILM 6) clustered with strains of zymodemes ILM 5 and 7 in all other analyses done here. Among other strains found to have the same ILM 6 profile, one was from Italy and another from Iran. MON 82 strains are known to occur sympatrically with MON 30 and MON 18 in Sudan. Furthermore, the cluster containing ILM 5, 6 and 7 did not appear close to the root of the *L. donovani* complex in any rooted robust tree of the *L. donovani* complex, and strains with these zymodemes were not reliably separated. I feel compelled to suggest that *L. archibaldi* does not warrant specific status and that this name should be discontinued, or that the definition of *L. archibaldi* be extended to include MON 30 and MON 18 as well. In the latter case, however, all five genetic groups of the *L. donovani* complex would have to be elevated to species, which is not practical.

Diagnosis and typing

Results from this work helped explain why it has been so difficult to find specific diagnostic tools for *L. infantum* in relation to *L. donovani*. Many characters are expected to be shared with other *L. donovani* groups, especially group E, but most importantly, if strains typed as MON 30 are considered as *L. infantum*, it will be impossible to find specific diagnostic tools. A rational search for such molecular tools, should include phylogenetic analyses followed by identification of autapomorphisms (or character states unique to a single taxon).

Characteristics and history of groups

Geographical correlations and genetic diversity levels varied in the *L. donovani* complex, depending on the genetic group. A strong geographical linkage was apparent in groups A (ILM 3) and B (Indian), and less in group E (ILM 10), whilst groups D (*L. infantum*) and C (ILM 5/6/7) are more cosmopolitan. A strong geographical linkage was also associated with low diversity in group B (Indian) but

not in group A (ILM 3 / Kenyan), which was one of the most genetically and phenetically diverse groups. Conversely, geographical dispersion was associated with high genetic diversity in group C (ILM 5/6/7) but not in group D (*L. infantum*), which was very conserved, except in the isoenzyme phenotypes. Group E (ILM 10) was mixed with low dispersion, low phenetic and *mspC* diversity, but with high RFLP and RAPD diversity.

Group A (mostly ILM 3) was polymorphic by all typing methods and had the most diverse *mspC* sequences, despite being restricted to Kenya and neighbouring regions of Ethiopia. This group may be ancient in the history of the *L. donovani* complex, but it appears relatively recent on *mspC* phylogenies and it is never near the roots of isoenzyme, RFLP or RAPD phylogenies. If indeed group A is recent or contemporary in origin to other *L. donovani* complex groups, two competing hypothesis are possible. Firstly, group A may have faced challenging and diverse habitats which would have thus exerted a positive selection for diverse genotypes. Secondly, the host/vector habitat of these strains may be more permissive than for other groups and thus letting through more diverse genotypes. Two Kenyan strains (MRC(L)3 and MRC74) usually clustered separately within the group or even outside the group. It is possible that it is an effect of longer culture or storage, but if the distinction is real, it could also suggest active speciation.

Not surprisingly, the Indian clade (group B) was very homogeneous, both phenetically (isoenzymes) and genetically (*mspC*, RFLP, RAPD). The Indian cluster was restricted to an isolated part of India and Indian *L. donovani* seem to be restricted to man. Indian kala-azar may have been a recent event, perhaps in a local adaptation to an exclusive human reservoir, or might have been the result of isolation of a small population (or a few founder strains). Both hypotheses would explain the low genetic diversity levels. This clade could have been spread with or within the Indian population but it did not. Perhaps local populations do not have migrating habits or perhaps Indian *L. donovani* are restricted to a local vector, the distribution of which would determine the geographical range of the parasite. Adaptation to a particular species of vector could also catalyse a speciation event and restrict genetic diversity. Indian strains other than MON 2 have been reported (MON 38) with distinct MDH, NH and PGD (Rioux *et al.* 1990). It would be interesting to determine if those strains belong to the Indian clade as determined in this project or if those are introductions from other genetic groups or even if those could be hybrids with strains from other groups.

Assuming that the Indian clade is the most recent, it was not possible to determine from the analyses done in this project which was their ultimate group of

origin. Phylogenies based on the *mspC* sequence pointed to group C (ILM 5/6/7), but isoenzyme based phylogenies favoured group A (ILM 3, Kenyan) and phylogenies of pooled RFLP, pooled RAPD/RFLP or RAPD did not favour group C in relation to group A. It is thus possible that the Indian clade originated from hybrid strains of clade A and C. Alternatively, the Indian clade may have diverged, instead, at the same time as those two groups from a more diverse population, having retained features that were also retained differently by the other clades, which would explain the apparent contradictory phylogenetic results.

Group C (ILM 5/6/7) strains can be found widespread in the Old World, mostly in the Mediterranean basin and the Horn of Africa, but more consistently in Sudan. As expected from its geographical but also host range, the diversity present in group C was high at all studied levels. This group may have benefited from a more plastic genome to conquer new habitats and find new ways of transmission, although group C may have been restricted to rodents, and limited by their ecology and distribution until recently when human population movements favoured dispersion of domestic rodents.

Remarkably, group D (*L. infantum / L. chagasi*) was simultaneously the most geographically dispersed and the second most genetically conserved group in the *L. donovani* complex. Some *L. infantum / L. chagasi* were not used in the overall analysis of the *L. donovani* complex, but they are widespread through tropical regions of the globe, from Western Europe to China, including the North of Africa, but also from South to North America. A somewhat similar process to that of group C may have happened with the *L. infantum* clade. *Leishmania infantum / L. chagasi* are mainly associated with canids, and particularly so with dogs. Not only may this association have produced the necessary isolation for emergence of a distinct clade, but also the association of dogs with man may have contributed to their faster and more effective spread. This group also seems to be very permissive in terms of vector and is thus capable of being transmitted by a large number of vectors present in new habitats. The most remarkable example is the colonisation of the American Continent by *L. infantum*, where it is transmitted by a different genus of sandfly.

Group E (ILM 10), instead was highly polymorphic in non-coding genomic regions, but had only one type of IEA profile and *mspC* sequence. This group is restricted to Ethiopia and a region of Saudi Arabia, close to Ethiopia, where these strains may have similar ecological niches, or where they may have differentiated recently. Interestingly, Saudi Arabian strains could not be distinguished from Ethiopian strains by any analysis, thus suggesting that one of the populations was

founded recently by migrants from the other (most likely Ethiopia, migrating to Saudi Arabia) or that the two populations have effective genetic exchange.

Whilst it is easy to envisage 'speciation' processes for both *L. infantum* / *L. chagasi* and Indian clades, it is more complicated to speculate on the differentiation processes which produced clades C, A and E, which are present in an almost cline of Sudan - Kenya - Ethiopia, respectively. It is possible that powerful geographical barriers contributed to their initial differentiation and that different adaptations to particular hosts or vectors may have maintained their geographical restrictions, with the exception of clade C (ILM 5/6/7) but probably in recent times only.

7.2.2. Analytical techniques

The gene tree (*mspC*) compared well with the more 'genome-wide' tree (RAPD, RFLP), especially regarding definition of clades.

The use of both RAPD and RFLP data, from several primers or enzymes and DNA regions, respectively, helped to overcome the limitations of gene phylogenies. A more realistic genomic tree was thus obtained. Although there is a possibility that the different typing methods were not compatible and therefore should not have been pooled, the trees generated have much more stable major groups than individual analyses, which, furthermore, were highly compatible to gene groups and IEA clusters. The methods used here thus seem complementary to each other. On the other hand, if RAPD and RFLP methods were analysed separately to build consensus trees (not shown), the main tree branches were not robust. Parsimony methods were the most amenable to consensus tree approachs, in that the main strain groups were identified, perhaps because of the large number of alternative trees produced in each analysis whilst clustering methods only produced one tree.

Distance and parsimony trees produced congruent topologies, which might be due to the large number of characters used. In such instances a cladistic approach has to weight multiple, often contradictory, synapomorphisms and may behave in a similar way to a numerical taxonomy approach.

As expected, the level of genetic diversity identified using RAPD and even RFLP was much larger than that of a functional gene sequence, such as *mspC*, or enzyme phenotype. RAPD/RFLP and *mspC* trees, however, were roughly compatible and differences could be explained by methodological constraints. It was interesting to observe that some groups with high internal diversity had only one *mspC* sequence. It is possible that whilst those groups are ancient and neutral mutations accumulate over time, the *mspC* sequence is maintained by selective pressure. In

such case it is an example of how a gene phylogeny can diverge from a genomic phylogeny.

It can be agreed that the most reliable data for phylogenies are DNA and protein sequence data, however, most data sets (as here) are limited to a single gene. RAPD and RFLP data, although not amenable to precise analysis (see above) can provide what may be considered as genome genetic data and thus generate phylogenies of organisms rather than that of genes. IEA data, although also providing information for several genes, is much more limited in size, less sensitive and data is much more difficult to analyse. Data from intergenic regions and anonymous regions (RAPD) may provide an estimate of neutral or nearly neutral evolution as opposed to the evolution of an enzyme coding gene. In that case, the *mspC* tree may not reflect the true history of the *L. donovani* complex. However, the major groups were congruent, and four of them could even be identified by specific *mspC* sequences.

It was apparent from comparison of UPGMA and neighbour-joining trees for pooled RAPD/RFLP data that the obtained distances were essentially ultrametric and compatible with the assumption of a molecular clock. Furthermore, by using pooled RAPD / RFLP data, all methods used for tree construction performed well in detecting the same genetic groups.

8. General Discussion

8.1. Evaluation of techniques

8.1.1. Typing methods

It became apparent from the results obtained in this project that genomic diversity and taxonomic studies should include several markers. These several markers should be of different natures, such as protein coding genes, intergenic regions and anonymous markers. Although most markers will correctly identify strain affinities, some problematic strains can only be satisfactorily typed with the aid of different methods. Determining which strains are problematic may not be easy from one analysis only and at least two reliable typing methods should be routinely employed.

- **Isoenzyme analysis**

Isoenzyme analysis, as used here, was useful for assignment of strains to groups because of the limited number of zymodemes found. Inclusion of more enzymes in the analysis, as for MON typing, increases the number of zymodemes but not enough to visualize the genetic groups in the *L. donovani* complex and may even further complicate analysis. Furthermore, the low level of discrimination obtained creates problems such as typing the Kenyan strain MRC74 as ILM 1, which was also the only zymodeme found in Indian *L. donovani*, and such as with the zymodeme ILM 8 that included both *L. infantum* and *L. donovani* strains. Therefore, phenotypic characters, like IEA, seem to be reliable for typing but not so much as phylogenetic markers, mainly because of their limited number and the uncertainty of distinguishing homologies from homoplasies.

- **RAPD**

RAPD was a much less reliable technique than IEA given the low reproducibility and its sensitivity to experimental conditions, but it was useful for detection of diversity at lower taxonomic levels. As such, RAPD was a sensitive method to analyse the relationship between *L. infantum* and *L. chagasi*, but also allowed detection of genetic clusters of *L. donovani*. Some strains' affinities, however, were not reliably determined. RAPD did not perform well at retrieving phylogenetic information between some groups of strains; tree topology changed according to the method used and bootstrap support was high only for some groups. Furthermore, the lack of reproducibility meant that different analyses could not be compared and restricted the number of strains or taxa that could be compared at any one time. RAPD is not a good method to detect heterozygosity unless homologous fragments are determined.

- **RFLPs**

Regarding the different RFLP typing methods, ITS RFLP seemed to be a good typing method, because there was less diversity within groups and groups could easily be identified by visual analysis of the RFLP patterns. However, some strains could not be reliably separated and group ITS-AB was not resolved, which included Indian strains (ILM 1) and Kenyan strains (ILM 3). Although ITS is a multicopy region, the low level, but robust, sequence diversity within groups may have been the reason for good separation of most groups.

By RFLP analysis of the mini-exon repeat unit it was not possible to identify reliably the genetic groups, although it could be useful to detect diversity within groups. The mini-exon is a very G-C rich region which limits the number of applicable enzymes. The mini-exon repeat unit had a high diversity level, in terms of fragment size and restriction sites, and it is full of nucleotide repeats, thus it may be a better marker for population genetics than for phylogenetics and taxonomy.

ITG/L did not produce reliable typing markers. One of the reasons may be that it is a multicopy target, but with a low diversity level. The PCR product also included part of the coding region, which may have complicated the analysis.

The best RFLP based typing method was ITG/CS because the main genetic groups were identified and strains were mostly correctly assigned when compared with the overall analysis of the *L. donovani* complex, except for strains Buck (I31) and Wangjie 1 (D26). This region is the only single locus analysed by RFLP which may be one of the reasons for its reliability, despite essentially corresponding to a gene phylogeny. ITG/CS was demonstrated to have a substantial degree of genetic diversity across *L. donovani* and within groups.

RFLPs can detect heterozygosity if the total fragment size is larger than the PCR product. In this project some strains were found to be heterozygous to several targets. However, no attempt was made to find putative parents to determine if those strains were hybrids or simple mutational heterozygotes.

- ***MspC* sequence**

Analysis of *mspC* DNA sequence arose as a useful method for analysis and typing of strains in the *L. donovani* complex. Not only were the same genetic clusters identified as by RFLP and RAPD analysis, but also because the resulting phylogeny has a sounder theoretical basis. Although this target may not be good for typing purposes because of the cost and labour associated with sequencing, some single nucleotide polymorphisms were identified which could be used in quick typing methods such as TaqMan, molecular beacons or nested-PCRs.

8.1.2. Phylogenetic analyses of the data

A problem with the phylogenetic analysis of RFLP is that the precise nature of all the data is not clear. The size of RFLP bands is dependent on both the size of the amplified fragment and on the presence of internal restriction sites. However, some polymorphisms also seemed to be due to simple divergence in size of the sequences, by insertions or by deletions, often in repetitive sequences. It was thus decided to treat each band equally as a separate character for analysis, as for RAPD data, irrespective of the origin of the size differences. In this project RAPD and RFLP data sets were pooled and the generated groups and phylogenies were congruent for bootstrap values higher than 90% to those obtained with individual data sets.

The best way of assessing genetic distances based on RAPD and RFLP data is debatable because the nature of variation is not known with precision in each case. Indeed, identification of homologous bands is uncertain without sequencing or some other form of sequence identification. There are models developed for restriction sites, which could not be determined with precision in this project and, thus, only restriction fragments were scored. The difficulty in determining restriction sites primarily arises from the complexity of the restriction profiles, which can have many causes, such as the long size of the fragments, the multiple copies of each target, the use of frequent cutting enzymes and the presence of insertions or deletions.

A phenetic approach was used here for phylogenetics, since a large number of characters was used, without many assumptions. Relatedness was simply determined by the proportion of fragments in common between two species (Jaccard distances) and analysed using clustering methods. A cladistic approach based on different parsimony methods was also applied, and the data were also analysed by maximum likelihood. Parsimony using large data sets can behave more like a phenetic method than a cladistic method, for it is necessary to conjugate often contradictory data, and parsimony analysis performed well in identifying genetic clusters when data were pooled and the method allowed for reversions to occur (Wagner parsimony). Genetic groups within the *L. donovani* complex were reliably identified by most phylogenetic methods.

Distance based methods are usually not universally supported for analysis of genetic data, because they do not have an intrinsic measure of reliability and information is lost in the process. In the present analysis, however, they were consistent throughout, with assumption of a molecular clock (UPGMA) or without such an assumption (neighbour-joining).

The analysis by several different methods to some extent allowed assessment of the reliability of phylogenies derived from the data. In fact, analyses of pooled data

were very similar whilst those of individual data (RFLPs, but also RAPD) were much less reliable. Either Wagner parsimony, UPGMA or neighbour-joining could be used in the future to analyse pooled RAPD or RFLP data, because group identification was congruent between methods and those groups were also supported by *mspC* sequence and, to some extent, isoenzyme analysis.

8.1.3. Identification of genetic clusters

Formation of genetic clusters was largely reproducible across the different techniques, except for the mini-exon and ITG/L (see above). The groups generated by RAPD and RFLP, as well as isoenzyme typing and *mspC* sequencing were consistent. The main groups of strains were:

- A - strains mainly from Kenya, mainly zymodeme ILM 3;
- B - strains from India, zymodeme ILM 1;
- C - strains from several locations, mainly from Sudan, zymodemes ILM 5, 6 and 7;
- D - strains from several locations, mainly zymodemes ILM 9 and 11;
- E - strains from Ethiopia and Saudi Arabia, zymodeme ILM 10.
- F - strain from China, zymodeme ILM 16.

All strains were affiliated with one group but some affinities were more difficult to determine, such as the case of strains Salti 4, Addis 142, Jeddah KA, Addis 164, Buck, MESH-17, Dora, MRC(L)3 and MRC74. Only strain Wangjie 1 was considered to be sufficiently distinct to constitute one group on its own (F).

Demes based on the present analyses were not assigned because of the large degree of diversity found, instead it was judged as more appropriate to find genetic families, which proved to be extremely interesting for typing, epidemiology and phylogenetic purposes. A possible exception might be *mspC* sequence for which only eight different types were found, and most groups only had one sequence type, which could be considered as 'DNAdemes': MSPC 1 (group E, ILM 10), MSPC 2 (group D, *L. infantum*), MSPC 3 (group C, ILM 5/6/7), MSPC 4 (group B, Indian), and MSPC 5-9 (group A, ILM 3).

Demes, such as zymodemes, schizodemes or 'RAPDemes', have little biological meaning unless included in a larger perspective, because they reflect information only at very restricted levels. The information conveyed can be contradictory and thus difficult to reconcile because integrated analyses are very rarely done. Furthermore, even if closely related strains often have very similar or identical profiles, if enough data are analysed, most strains could be differentiated. The best practice, then, would be to use two or three carefully chosen typing techniques; two if they agree and a third if they disagree.

From the data presented in this project, it was found that ITS and gp63 ITG/CS RFLP may provide enough reliable and reproducible data for typing to genetic family. RAPDs could be a good typing method, but require the inclusion of several reference strains for accurate typing and are very difficult to apply correctly in practice. Mini-exon and gp63 ITG/L should only be used if more sensitive typing is required, such as for tracking of epidemics and population genetics. The DNA sequence of *mspC* provides more clear cut typing but may be difficult to apply in most laboratories.

8.2. Genetic characterization of Portuguese *Leishmania infantum* strains

As verified by IEA, most Portuguese *L. infantum* strains were genetically very similar by RAPD, in comparison with *L. infantum* strains from other areas. Most strains analysed by RAPD had been isolated from the canine reservoir, four from the sandfly vector and one from a fox. Most strains were ILM 9 (putative MON 1 or LON 49), but two of them (IMT 171-2), which were isolated from the vector in the North, were MON 24 (ILM 11) and these clustered separately in RAPD trees. It was not clear whether these strains could have originated from local strains, as the neighbour joining tree suggests, or they are outside the local clade, as the single linkage tree indicates (Fig. 22). All other strains were very closely related and only a small cluster of three strains, two isolated from the vector and one from a dog in the North and in the same campaign, could be distinguished from the rest.

Unfortunately, the analysis was limited both in number of strains and number of RAPD primers, but most internal branches were very short and it is possible that the population structure of Portuguese *L. infantum* can not be resolved by RAPD analysis. Perhaps the use of more defined genetic markers, such as microsatellites, can elucidate the epidemiology of *L. infantum* in Portugal. It would also be beneficial to have an Iberian analysis, rather than a Portuguese restricted study, in order to investigate the possibility of migrations.

Two *L. donovani* strains which had been isolated in Portugal were also studied here. These strains were very similar and belonged to the same genetic group (C, ILM 5/6/7). One of the strains had been isolated from an HIV+ patient (IMT 180) and the other (IMT 188) from *S. minuta*, putative vector of *L. (Sauroleishmania)*. It had been assumed that the HIV+ patient had been infected abroad or intravenously, since no *L. donovani* had been previously known in Portugal. However, the isolation of the second *L. donovani* strain from a lizard biting vector, suggests that either the parasite is currently being transmitted in Portugal, possibly from HIV+ patients, initially or that there was an unidentified *L. donovani* cycle in Portugal, from which the HIV+ patient

was accidentally infected. These findings, and the increasing number of unusual *Leishmania* strains isolated from HIV+ patients, stress the importance of a detailed knowledge of VL epidemiology in endemic regions, including unforeseen hosts and vectors.

8.3. Synonymy of *Leishmania infantum* and *Leishmania chagasi*

8.3.2.The mysterious origin of *Leishmania chagasi,* or a return home

The only known agent of visceral leishmaniasis in the Americas is named *L. chagasi*. However, the other species of the *L. donovani* complex are found in the Old World. Several enzymatic and genetic methods have indicated that *L. infantum* and *L. chagasi* were very close related and some authors propose synonymy (Rioux *et al.* 1990; Grimaldi & Tesh, 1993). Some other authors do not agree (Shaw 1994) and have used minor phenotypic and genotypic differences (Decker-Jackson & Tang, 1982; Santoro *et al.* 1984; Palatnik *et al.* 1990; Ellis & Crampton, 1991) to justify separation into two distinct species. In the case of two distinct species *L. chagasi* is said to have been present for a long time on the American continent. On the contrary, most studies of these species have shown a low level of diversity between *L. infantum* and *L. chagasi*, suggesting that separation was quite recent and probably occurred by importation with infected dogs accompanying Spanish and Portuguese colonists (Killick-Kendrick 1980; Momen *et al.* 1993).

Ecological and epidemiological characteristics of the parasite have been interpreted as evidence for the prolonged presence of *L. chagasi* in the Americas (Lainson & Shaw 1987; Travi *et al.* 1998). It has been said that the severity of disease seen in dogs would prevent any infected animal surviving a long voyage to the New World. The incubation period, however, can be very long (Rioux *et al.* 1979; Vexenat personal communication) and subclinical infection, with infectivity for sandflies, is common (Vexenat *et al.* 1993, Molina *et al.* 1994). The finding of healthy infected foxes in the New Word might also suggest an ancient association. However, the potential of New World foxes to be a reservoir of *L. infantum* was high since they are closer to Old World canids than to Old World foxes. Furthermore, severity of disease does not necessarily correlate with duration of host-parasite association (Combes 1997; Lipsitch & Moxon 1997), and infection in dogs is also frequently asymptomatic. The existence of non-canid reservoir hosts in the New World has also been used as an argument for the ancient presence of *L. chagasi*, but infection seems to be more dependent on feeding habits of the vector and not to require special adaptation. *Leishmania infantum* is known to infect rodents in the wild in

Europe (*Rattus rattus*) and also in the laboratory (*Mus spp.* and *Mesocricetus auratus*). The adaptation of the parasite to at least two vectors in the New World would seem very difficult in a short time, and yet *Lu. longipalpis* was shown to be as susceptible to *L. infantum* infection as *P. ariasi* (a vector in Europe), New World sandflies have less parasite specificity than OW sandflies (Killick-Kendrick *et al.*1980) and *L. infantum* can be transmitted by a variety of different *Phlebotomus spp.* in the Old World. The wide geographical range of *L. chagasi* in the New World may not have required a long history because both Spanish and Portuguese may have introduced the parasite independently. Movement of populations, facilitated by trade, common or similar languages and cultures, not forgetting that those two countries were politically united for about 60 years in the 17th century, may have been another factor facilitating a rapid spread of the organism. The lack of genetic diversity and remarkable overall similarity with *L. infantum* are not compatible with the introduction of *L. chagasi* to south America 2-3 My ago, simultaneous with the arrival of wild canids in the New World (Lainson *et al.* 1987), if the estimate of less than 2My for the separation of *L. infantum* from *L. donovani* (Moreno *et al.* 1984) is correct.

The amount of genetic and enzymatic data assembled to the present time and in this project strongly suggest that *L. infantum* and *L. chagasi* are synonymous. The many arguments against synonymy, as exposed above, seem to be invalid, although they are still being put forward (Travi *et al.* 1998), despite the amount of evidence against them. Here, not one solid argument could be found for the prolonged presence of *L. chagasi* in America, and therefore, all present data are compatible with introduction after the arrival of European colonists. We believe that *L chagasi* are *L. infantum* parasites that returned to the place where *Leishmania* most likely originated, that is South America. As explained above it is not surprising as *L. infantum* was clearly pre-adapted to colonise a different continent in a short time.

It is possible that *L. chagasi* is undergoing speciation by allopatry, but this process may be complicated by re-introductions and migrations to and from the Old World.

8.4. Genetic diversity in the *Leishmania donovani* complex

8.4.1. Clustering of strains into genetic groups

Through the use of different techniques it was possible to define in this project at least five genetic groups within the *L. donovani* complex, as shown in Table XXXII, and as described in 8.4.2. Any relationships with the year of isolation were not

apparent from the analysis, although there were strong geographical and isoenzyme associations.

Most groups were robust and easily defined, but assignment of some strains was difficult. If a number of methods was considered the affinities of problematic strains became apparent.

- strain Addis 142 (D15) was placed in group A by IEA, RFLP and RAPD/RFLP, but clustered with group E in RAPD.
- strains MRC(L)3 (D2) and MRC74 (D3) belonged to group A by every technique, except that they clustered independently by ITG/L and with group B by ITS RFLP.
- strain Salti 4 (D17) clearly belonged to group C by every technique, except by RAPD, in which it was placed near the root.
- strains MESH-17 (D5) and Dora (D24), belonged to group C but clustered in group F(B) in RAPD.
- strain Buck (I31) is *L. infantum* (group D) but did not cluster reliably by ITG/CS and ITG/L.
- strain Jeddah KA (D32) clustered with group E in all techniques, except that by RAPD it was placed with group C.
- strain Addis 164 (D29) clustered with group E in RFLP and RAPD/RFLP trees, but was closer to group A in RAPD.

Other strains were much more problematic. Strain WR 341 (C1) was not included in most analyses because its identity was not certain and thus its characterization would not be relevant. Affinities of strain Sukkar 2 (D23) were impossible to determine as it was so different from any other, and most analyses were not even scored, although it was included in most analyses.

8.4.2. Description of genetic groups

- **Group A**

Strains in group A had been isolated from Kenya, mainly, but also from neighbouring regions of Ethiopia. The group could be defined by zymodeme ILM 3, although two strains (MRC74 and D2) had zymodemes ILM 1 and 8, respectively. The *mspC* sequence was polymorphic, but RAPD and ITG/CS RFLP produced a defined cluster. Group A, however, could not be distinguished from group B upon ITS RFLP. Strains MRC(L)3 and MRC74, from Kenya, seemed to be special cases within group A, because, they usually formed a sub-group. It is possible that a definite sub-group could have emerged if more related strains were included in the analysis. It is also possible, but less likely, that the more awkward genotypes might be due to long term culture of these strains.

Table XXXII - Group assignment of *L. donovani* complex strains.

Code	WHO code	Zym			*MspC*	RAPD	ITS	ITG/ /CS	RFLP	Group
		LON	MON	ILM						
D2	MHOM/KE/1967/MRC(L)3	-	-	3	A1	A	AB	A	A	
D3	MHOM/KE/1973/MRC74	51	-	1	A2	A	AB	A	A	A
D15	MHOM/ET/1984/Addis 142	-	-	3	-	E	AB	A	A	(kenya)
D16	MHOM/KE/1980/Ndandu 4A	44	-	3	A3	(A)	AB	A^h	A	ILM 3
D21	**IMAR/KE/1962/LRC-L57**	44	37	3	A4	A	AB	A^h	A	
D25	MCAN/KE/0000/D2	45	-	8	A4	A	AB	A^h	A	Kenya/
D27	MHOM/ET/0000/Ayele 8	56	-	3	A5	A	AB	A^h	A	/Ethiopia
D30	MHOM/KE/0000/Neal R1	56	-	3	-	A	AB	A^h	A	
D31	MHOM/KE/1975/Mutinga H9	56	32	3	A1	-	AB	A^h	-	
D	**MHOM/IN/1980/DD8†**	41	2	1	B1	B	AB	B	B	
D4	MHOM/IN/1982/Patna 1†	41	-	1	B	B	AB	B	B	B
D6	MHOM/IN/1977/Chowd X†	-	-	1	B	B	AB	-	-	(india)
D7	MHOM/IN/1979/STL1-79†	-	-	1	B	B	AB	B	B	ILM 1
D8	MHOM/IN/1982/Nandi 1†	41	-	1	-	-	AB	B	B	India
D1	MHOM/ET/1967/HU3(LV9)	46	18	5	C	G(C)	AB	C	C	
D5	MMER/IR/1996/Mesh 17	50	-	6	-	F(B)	C	C	C	C
D17	MHOM/LB/1984/Salti 4	-	-	5	C	U	C	C	C	(sudan)
D18	MHOM/SD/0000/Khartoum	46	18	5	C	-	C	C	C	ILM
D19	MHOM/SD/1985/A22	-	-	7	C	G(C)	C	C	C	5/6/7
D22	MARV/SD/1962/LRC-L64	48	-	7	-	U(C)	C	C	-	
D24	MCAN/IT/1976/Dora	50	-	6	C	F(B)	C	C	C	Sudan
D28	MHOM/ET/1972/Gebre 1*	50	82	6	C	U(C)	C	C	C	
D33	**MHOM/SD/1982/Gilani**	48	30	7	C	C	C	C	C	*L. archi-*
D34	MHOM/PT/1992/IMT 180	-	18	5	C	C	C	C	C	*baldi?*
D35	ISER/PT/1993/IMT 188	-	-	5	C	C	C	C	-	
I	**MHOM/TU/1980/IPT 1**	49	1	9	D	D	D	D	D	
C	MHOM/BR/1974/PP75[a]			9	D	D	D	D	D	D
I4	MHOM/ES/1987/Lombardi	-	-	11	D	D	D	-	-	(infantile)
I16	MHOM/CN/1980/Strain A	-	34	9	D	D	D	D	D	ILM 9
I17	MHOM/CY/1963/L53	-	-	12	D	D	D	D	D	(ILM 11)
I25	IARI/PT/1989/IMT 171	-	24	11	D	D	D	-	-	*L. infan-*
I26	IARI/PT/1989/IMT 172	-	24	11	D	D	D	D	D	*tum*
I31	MHOM/MT/1985/Buck	49	78	11	D	D	D	A	-	
D9	MHOM/SA/1987/VL23	-	-	10	-	-	E^h	E	E	
D10	MHOM/SA/1987/VL29	-	-	10	E	E	E^h	E	E	E
D11	MHOM/SA/1987/VL6	-	-	10	-	E	E^h	E	E	(red sea)
D12	MHOM/ET/0000/Ayele 5	52	-	10	E	E	E	E	E	ILM 10
D13	**MHOM/ET/0000/Hussen**	42	-	10	E	E	E^h	E	E	
D14	MHOM/ET/1982/Bekele	42	-	10	-	E	E	E	E	Ethiopia/
D20	MHOM/SD/1987/UGX-marrow	-	-	10	-	E	E^h	E	E	/Saudi Arabia
D29	MHOM/ET/1984/Addis 164	-	83	10	E	U(A)	E^h	E	E	
D32	MHOM/SA/1981/Jeddah KA	42	31	10	-	C	E^h	E	E	
D26	**MHOM/CN/0000/Wangjie 1**	-	35	16	-	U(A)	E^h	F	(E)	F? (china)
C1	MHOM/PA/1980/WR341	-	-	14	-	-	E?	-	-	-
D23	MCAN/IQ/1981/Sukkar 2	43	-	13	I	U	-	-	-	-

In parenthesis are external affiliations to groups. U is ungrouped and in parenthesis, underlined, are majority consensus groups. * - *L. archibaldi*. In bold are strains suggested here as reference strains for their group. [h] are putative heterozygotes. [a] *L. infantum* (*L. chagasi*) alternative reference strain.

- **Group B**

Group B is characterised by zymodeme ILM 1 and only contains strains from India. This group had only limited genetic diversity and formed a robust cluster, which, however, could not be separated from group A upon ITS RFLP.

- **Group C**

Strains included in group C were typed as three zymodemes, ILM 5, 6 and 7. Although most of the strains, and according to the literature, had been isolated in Sudan, this group is dispersed over a number of countries and not only in Africa. The group was not robust upon RAPD but was consistent upon other techniques.

According to this analysis there is no support for a specific status or sub-specific status of *L. archibaldi* (ILM 6). The type strain of this group clustered with group C, which included zymodeme variants ILM 5, 6 and 7 and is extremely cosmopolitan. This group had been placed in the root of the *L. donovani* complex by Rioux *et al* (1990) and from which three species (*L. infantum*, *L. donovani* and *L. archibaldi)* were defined. In the present IEA, however, zymodeme MON 82 (ILM 6) emerged as a putative hybrid of ILM5 and ILM7 and in another analysis the group was reliably placed far from the root of the complex. Indeed, as Cupolillo *et al.* (1998) had already pointed out, descriptions of species or groups should not be based on a single character.

- **Group D**

Strains in group D were included in the species *L. infantum* and *L. chagasi.* These strains formed a robust clade with low genetic diversity, but some phenotypic diversity (ILM 8, 11, 12). It was here demonstrated that the two species are synonymous, and thus this group is the most geographically diverse of the *L. donovani* complex.

- **Group E**

Strains in group E were restricted to Ethiopia and Saudi Arabia, although strain UGX-marrow was from Sudan. These strains were all ILM 10 and also formed a robust clade. Genetic diversity was high, but *mspC* and isoenzymes were very conserved.

- **Group F**

A possible sixth group might include Far-East strains such as strain Wangjie 1 (from China). This strain was difficult to place reliably in this analysis, which may be due to being the only one of its kind used in this project.

8.4.3. Biological significance of the groups

It is not clear whether the five clades described in this project have any biological or clinical significance. Group B (Indian strains) had good geographical, if not clinical and epidemiological, fit. Group A also fitted with Kenya and also Ethiopia, and group E with Ethiopia and Saudi Arabia. Unfortunately, the reservoirs for these latter groups have not yet been established. Group D (*L. infantum*) was the only group with recognised clinical and epidemiological significance.

PKDL is commonly associated with Indian strains, whilst the *L. infantum* clade is not known for it. PKDL is also present but rare in Africa, although it was frequent in a recent epidemic in Sudan (Zijlstra *et al.*, 1994) and it would be interesting to investigate the association of the different groups with PKDL. However, it is possible that PKDL may be a host dependent condition.

8.4.4. Clonality versus sexuality

Although *Leishmania* are considered to have a clonal population structure, some putative hybrids have been found which suggest the possibility of genetic recombination in natural populations. No hybrids have been described which involved *L. donovani* complex strains, but some clues were found here which suggested hybridization or recombination events within the *L. donovani* complex. No strain, however, was found to have a definite hybrid profile in all methods employed.

Zymodeme ILM 6 identified in group C ('*L. archibaldi*') is heterozygous for ASAT, and could be a hybrid between zymodemes ILM 5 and 7. Zymodemes ILM 5, 6 and 7 may form or have formed a population with recombination, probably still active in some areas of Sudan. Alternatively, an initial 5 or 7 phenotype may have suffered a mutation in one of the ASAT alleles, thus forming a heterozygote (ILM 6), which by meiosis or unequal mitosis could have produced the third homozygote. However, the existence of another ASAT heterozygote, as discussed in Chapter 3, indicated that recombination was the most likely scenario.

Strains in group B (Indian strains) were clearly polymorphic for mini-exon, with two alleles of different sizes, and some strains also appeared to be heterozygous upon RFLP analysis. The total size of the RFLP fragments was larger than the PCR product or significantly larger than or most other strains, which suggests the presence of different alleles. Some 'heterozygous profiles' were more common in certain groups (Table XXXII): in group A (ILM 3), ITG/CS and ITG/L (D15, D16, D21, D25, D27); in group E (ILM 10, Ethiopia/ Saudi Arabia), ITS and mini-exon (D9-D14); in group C, mini-exon (D1, D5, D18, D19, D28, D33-D35). Not all strains in one group, neither strains in group E (*L. infantum*), nor each mini-exon PCR product in group B

(Indian strains), however, had heterozygous profiles. It was not possible in this project to determine whether the different alleles had arisen independently, if the different alleles were at the same loci or if heterozygous strains were hybrids.

The occurrence of genetic recombination in local *L. donovani* complex populations would explain why genetic groups have been so clearly identified instead of a multitude of different clades. If the population structure of the *L. donovani* complex is mainly clonal, a multitude of clades would be expected. Thus, most ancient clades would have to have been extinguished, and each clonal group would descend from a single strain. The possibility of recombination in the *L. donovani* complex, thus, deserves to be investigated. *MspC* markers could be used in laboratory experiments and microsatellites would be invaluable in research on local populations.

8.4.5. Practical significance of the groups

From this project, the phylogenetic conclusions were not considered to be as important as the production of a working basis for analysis of the *L. donovani* complex. Using mainly ITG/CS and ITS RFLP or sequence analysis of *mspC*, but also RAPDs and even IEA, strains can be assigned to one of the groups.

It is never desirable to increase the number of *Leishmania* species, least of all within a conserved complex such as the *L. donovani*. However, if the *L. donovani* complex genetic groups are found to have biological, clinical or epidemiological significance it may be useful to find working designations. One of the groups already had the specific designation of *L. infantum* or *L. donovani infantum*, depending on the authors and could be named *L. donovani* (infantile), but there is no designation for any of the other groups. The designation *L. archibaldi* has been used for zymodeme MON 82 (ILM 6, LON 50), but that could be extended to include zymodemes MON 30 (ILM 7, LON 48) and MON 18 (ILM 5, LON 46) which were found here to belong to the same genetic group C, although it would be better a new designation, *L. donovani* (sudan). The designation *L. donovani s.s.*, although not fully correct, could also be applied to designate the Indian *L. donovani*, as has been in use by LSHTM researchers, but *L. donovani* (india) would be preferable. Designations for groups A (ILM 3), E (ILM 10) and F (ILM 16) are not available. Because the remaining groups seem to have clear geographical associations, the designation *L. donovani* (china) could be used for group F (Wangjie 1), from China, and the designation *L. donovani* (kenya) could be used for group A (ILM 3), from Kenya / Ethiopia. The designation *L.*

donovani (red sea) could be used for group E (ILM 10), from Ethiopia / Saudi Arabia. These designations will be used from here on.

Correct group identification depends not only on the method, but also on appropriate reference strains, especially when applying methods which require comparison of bands, such as IEA and RAPD but also RFLPs. It has been suggested that Indian strain DD8 is not a good *L. donovani* reference strain because it is polymorphic for the mini-exon. In the present project, this feature was confirmed and seemed to be present in most, if not all, Indian strains, but not in any other groups. Therefore, DD8 could still be used as a reference strain for Indian *L. donovani*, but should be used alongside reference strains for other *L. donovani* groups if a more detailed analysis is necessary. Strains that were systematically and reliably placed in the correct groups by all techniques used and have been used by other researchers would make good references. HU3 and IPT 1 have been used as reference strains and could still be used for *L. donovani* (sudan) and *L. donovani (infantile)*, respectively. Strain HU3 was not correctly placed by ITS, however, and strain Gilani might be a better reference for *L. donovani* (sudan). *Leishmania donovani* (kenya) and *L. donovani* (red sea) did not include any traditionally used reference strains. For *L. donovani* (kenya), strain LRC-L57 fills these conditions, whilst for *L. donovani* (red sea) strain Hussen is the best candidate.

The definition of genetic groups within the *L. donovani* complex may be extremely important for establishment of better epidemiological models. Whilst finding the main host of the *L. donovani* (infantile) and *L. donovani* (india) clades was easy, because they are very specialised, more generalist and ancient clades may survive in several hosts, depending on historical conditions if transmission is maintained by the vector. The finding of the responsible host for African kala-azar in Sudan - Ethiopia - Kenya may be difficult for different reasons. Generalist groups, such as *L. donovani* (sudan) may be able to colonise a number of hosts, whilst *L. donovani* (red sea) may have a much more human dependency.

8.4.6. Phylogeny of the *Leishmania donovani* complex genetic groups

One of the main questions regarding the phylogeny of the *L. donovani* comples is how did the genetic groups emerge. The geographical association of some groups seems to favour a scenario with isolation of small populations, but other factors, such as host, vector, and climate, may have been important for some groups. It is therefore apparent that several effects may have acted on the evolution of the *L. donovani* complex, depending on local and historical events. Factors involved in evolution of each group have been discussed in Chapter 7.

The other main question is how did the groups evolve in relation to each other. From the data gathered here and from the literature, two different models of evolution of the *L. donovani* complex are described: an integrated model, which was deduced from *mspC*, RAPD and RFLP phylogenies, obtained in this project, and a zymodeme model, which was based on the isoenzyme phylogeny of Rioux *et al.* (1990). Whilst the first model deals with evolution of genetic groups, the second deals with evolution of zymodemes, but the genetic groups can be conjugated with the isoenzyme phylogeny because they were found to correlate strongly with zymodeme types or groups. The two models, however, differ in rooting the complex and thus the two alternative scenarios are proposed (Fig. 53 and 54). Unfortunately, many hosts for African *L. donovani* are still poorly or not known and it is thus difficult to speculate on their importance for 'speciation' within the complex.

The zymodeme model in Fig. 53 places three main genetic pools in Africa, which may have differentiated from an ancestral population in a Sudan - Ethiopia - Kenya cline, but had been rooted to the *L. donovani* (sudan) by Rioux *et al.* (1990). According to this model, *L. donovani* (kenya) would have differentiated from *L. donovani* (sudan), perhaps by association with different vectors, subgenus *Synphlebotomus* and subgenus *Larroussius* respectively. However, the separation of the two clades could have been only coincidental with different distributions of the vector and be due to geographical isolation or association with different hosts. From *L. donovani* (kenya) would have originated the clade which became established in India and is now recognised as *L. donovani* (india), which would also explain the profile of a *L. donovani* (kenya) strain only distinguishable by small differences in the MDH profile from ILM 1. Zymodeme ILM 1 is not found anywhere else but India and the vector belongs to a different sub-genus. The ancestor of *L. donovani* (red sea) would have become isolated in or migrated to Ethiopia, also from *L. donovani* (kenya) strains. This group may have been introduced recently into neighbouring Saudi Arabia, possibly by rodents transported by ships or by infected humans. The origin of *L. donovani* (china) is more obscure as it could be a descendent of either *L. donovani* (kenya) or *L. donovani* (red sea). The *L. donovani* (china) group could be a remnant of related strains which could have spread early through Asia (possibly before the Indian group). Events such as a glaciation, or simply geographical distance, might have restricted its distribution to East China.

On the other main branch of the tree, are the *L. donovani* (infantile). According to the isoenzyme model, they would have branched from *L. donovani* (sudan) strains. The main hosts of the infantum group are canids, the adaptation to which may have contributed to speciation of this group. *Leishmania donovani* (sudan) have also been

found in carnivores as well as rodents and it is a very likely ancestor for more specialised groups. The two clades are transmitted by the same vector subgenus, although *L. donovani* (infantile) is known to be very versatile in that respect for it is also transmitted by *Lutzomyia* in the New World. These two clades are both generalists, because strains of the ILM 5, 6, 7 clade (Group C) were also found in Italy in rats and Iran in rodents, perhaps in recent introductions driven by human action. *Leishmania donovani* (infantile) is the most successful clade of *L. donovani*, for it spread across 4 continents. The extreme success of *L. donovani* (infantile) may be due to a close association with the genus *Canis*, which domestication may have allowed not only a greater geographical diversity as an increased probability of transmission in permanent settlements.

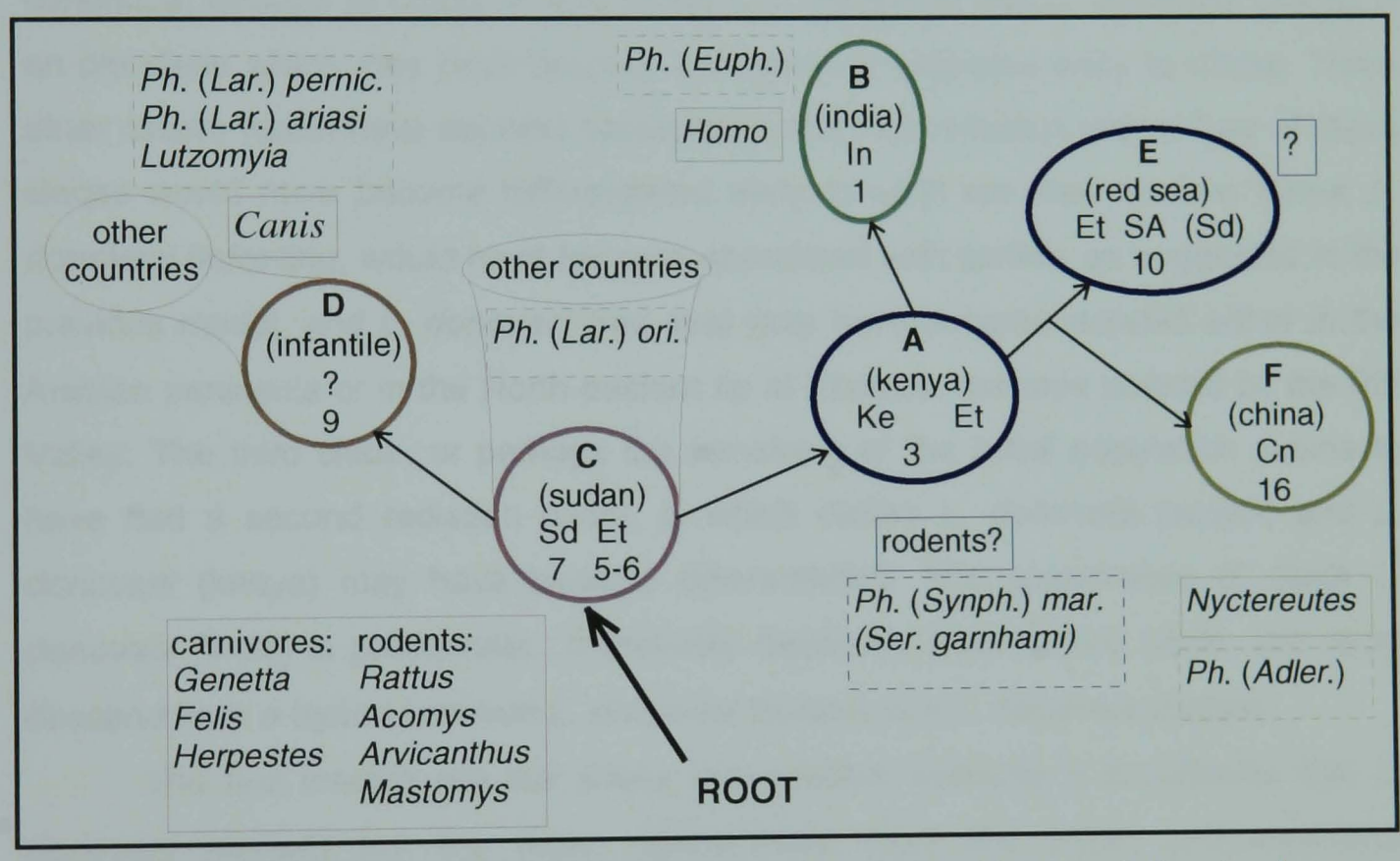


Figure 53 - Schematic representation of the speculative history of the *L. donovani* complex based on the isoenzyme phylogeny by Rioux *et al.* (1990), clarified by the genetic groups identified here. Genetic pools are represented by thick circles, with the country and zymodeme (below). Hypothetical reservoirs are in boxes and vectors in dotted boxes. Key for countries: Cn is China, Et is Ethiopia, Ke is Kenya, In is India, SA is Saudi Arabia and Sd is Sudan. Key for vector: *Ph.* is *Phlebotomus*, *Lutz.* is *Lutzomyia*, *Ser.* is *Sergentomyia*, *Adler.* is for *Adlerius*, *Lar.* is for *Larroussius*, *Syn.* is *Synphlebotomus*, *mar.* is *martini*, *ori.* is *orientalis*, *per.* is *perniciosus*. Although all groups are known to infect humans, these are the exclusive reservoir only for ILM 1.

The alternative representation of the evolution of the *L. donovani* complex is an integrated model (Fig. 54), based mostly on *mspC* phylogenies, but also RFLP and RAPD phylogenies, and has a very different topology from the above, although particular speciation events can be similar. This model is based on an initial population of visceralizing strains, and suggests two radiative events, one at the origin of the main clades and the other within one of the clades. It is tempting to speculate that the initial population may have been constituted by generalist strains, or by a mixture of strains with different specificities. This initial population may have existed either in Africa (Sudan - Kenya) or in Asia (Near - Middle East) and might have been a small population restricted to an area in Northeast Africa or the Near East, probably with a high recombination frequency, or may have been spread in a continuum through to China. In this model it is suggested that *L. donovani* (china) is an old clade which may have become isolated in or migrated early to China. Three other clades would have become separated in the first radiation event. Two of these clades would have become differentiated early to what are their modern forms; *L. donovani* (infantile), would have become associated with canids, as suggested in the previous model, and *L. donovani* (red sea) may have become isolated either in the Arabian peninsula or in the North-eastern tip of Ethiopia, perhaps isolated by the Rift Valley. The third clade, or perhaps the remaining of the initial population seems to have had a second radiation event, in which clades *L. donovani* (sudan) and *L. donovani* (kenya) may have become differentiated. The appearance of clade *L. donovani* (india) is problematic. It probably descends from a third clade, but could descend from a hybrid between *L. donovani* (sudan) and *L. donovani* (kenya).

The two models are not totally incompatible because it is possible that *L. donovani* (sudan) are the direct descendants from the initial undifferentiated population. This clade could have maintained its host generalist features and possibly its zymodeme phenotypes, although suffering evolutionary change, mainly in non-coding DNA sequences such as the intergenic sequences analysed by RFLP, but also in genes such as *mspC*. This scenario would explain the isoenzyme rooting of the zymodeme model. The two models, however, differ in the position of groups *L. donovani* (Red Sea) and *L. donovani* (infantile), although the integrated model is probably more correct because it is based on different methods. The *L. donovani* (Red Sea) and *L. donovani* (infantile) groups are distantly related and may appear to descend from different groups in the zymodeme model because they may have become individualised from distinct segments or strains of the primordial population. When the different segments in the ancestral population later 'speciated', their

isoenzyme profiles may have changed little and thus the group would appear closer to other, in fact, distantly related strains or zymodemes.

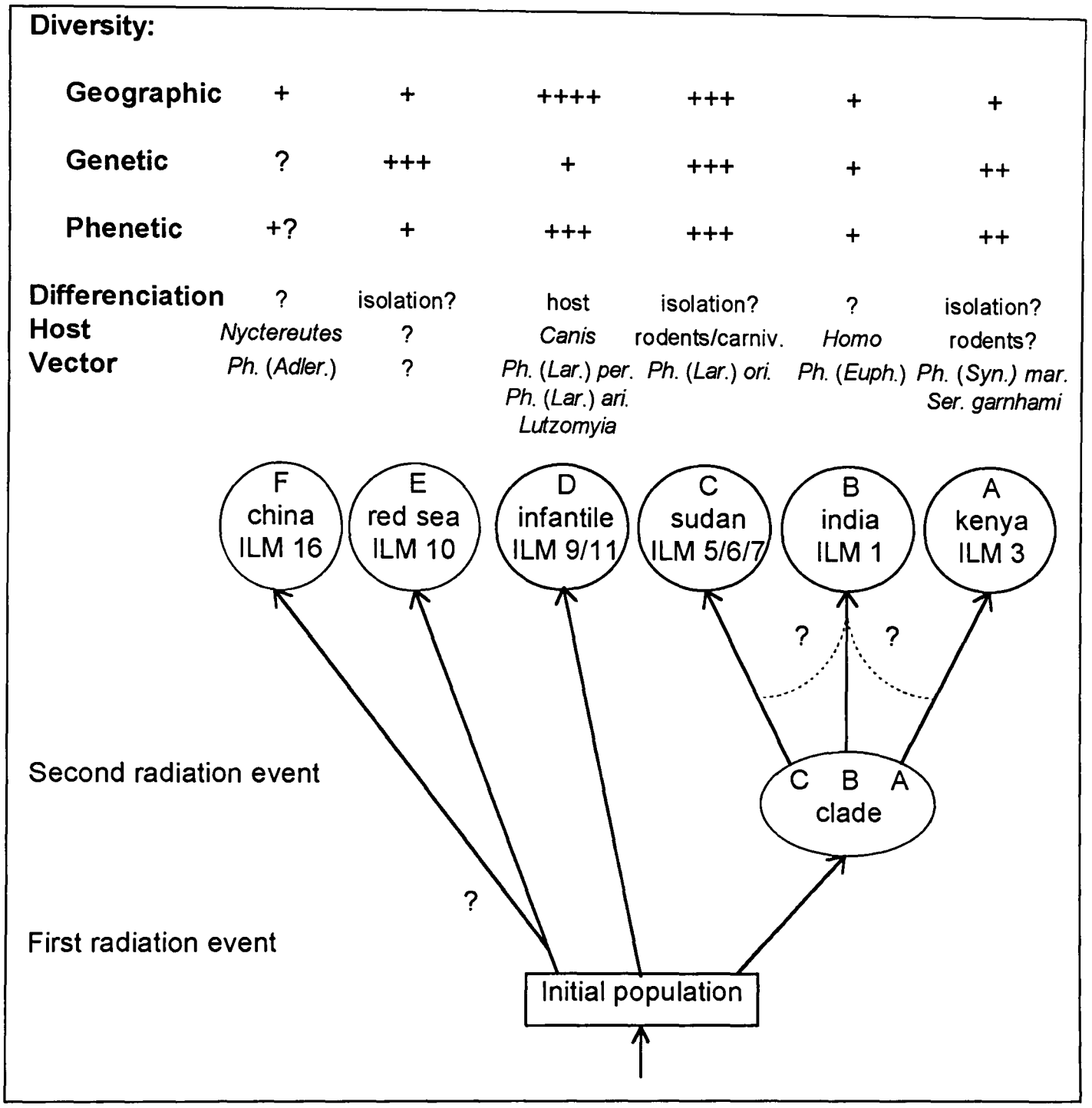


Figure 54 - Speculative alternative to the history of the *L. donovani* complex based on *mspC* and RAPD / RFLP data. Dashed lines are highly speculative. Scale of diversity is + (low) to ++++ (very high). Key for vector: *Ph.* is *Phlebotomus*, *Lutz.* is *Lutzomyia*, *Adler.* is for *Adlerius*, *Lar.* is for *Larroussius*, *Ser.* is for *Sergentomyia*, *Syn.* is *Synphlebotomus*, *mar.* is *martini*, *ori.* is *orientalis*, *per.* is *perniciosus*.

The *L. donovani* (infantile) and *L. donovani* (india) clades share a low degree of genetic diversity and a strong association with humans, which may indicate a relatively recent origin, probably in close association with human action: the *L. donovani* (infantile) clade indirectly by domestication of dogs and the Indian clade directly by development of a human exclusive cycle. However, and especially in the case of *L. donovani* (infantile) this may not be entirely true. Whilst it is likely that

L. donovani (india) are only as old as established human populations in India, *L. donovani* (infantile) can be more diverse than *L. donovani* (india) and may seem more recent due to a genetic bottle-neck at some point during their history. This might be the preferential survival of *Canis familiaris* strains or reduction to a small population during a glaciation period. The *L. donovani* (infantile) appear as recent as wolves / dogs (1-0.1Mya, Pleistocene) (Vila *et al.*, 1997; Wayne *et al.*, 1993) but are probably older, even as old as wolf-like canids (9Mya, Miocene) (Wayne *et al.*, 1993). One of the apparently old groups, *L. donovani* (chinensis), is associated with racoon dogs, which have originated in the Miocene (Wayne *et al.*, 1993). Such dating would put the date of the first *L. donovani* complex radiation in the Miocene. To resolve this issue it would be necessary to investigate the presence of *L. donovani* (infantile) in wolves and jackals in wild populations throughout the Old World.

In order to test these hypotheses it would be necessary to study strains from Central Asia, as well as China to test the existence of a cline, and of different hosts. Robust rooting of the *L. donovani* complex and a molecular clock are imperative to test whether *L. donovani* started differentiating in Africa or Asia .

8.4.7. Speculations on the origin of the *Leishmania donovani* complex

Some published phylogenies of *Leishmania* show the *L. donovani* complex as a separate clade from OW CL complexes (*L. major*, *L. aethiopica*, *L. tropica*) (Piarroux *et al.*, 1995; Croan *et al.*, 1997) or even as separately from all CL complexes (Schonian *et al.*, 1996). Most available phylogenies, however, do not resolve species of the sub-genus *Leishmania*. The *L. donovani* complex as a whole is very versatile because it has varied hosts, from rodents to carnivores, and vectors (subgenus *Larroussius*, *Euphlebotomus*, *Synphlebotomus* and *Adlerius*) and it is thus difficult to establish which would be the ancestral condition. One intrinsic factor, however, might help explain the differentiation of a VL clade. Visceralizing *Leishmania* are known to tolerate higher temperatures than cutaneous *Leishmania* (Berman and Neva, 1981; Callahan, *et al.*, 1996; Fehniger, *et al.*, 1990). The ancestors of visceralizing strains may have become adapted to higher temperatures than cutaneous strains, thus being able to survive in a wider range of habitats, namely, internal organs of hosts and, perhaps, survive more extreme temperatures in the vector. This suggestion would have to be tested in terms of *L. donovani* complex genetics and ecology of hosts and vectors.

The *L. donovani* complex seems to be an ancient branch of the OW *Leishmania* and the cutaneous *Leishmania* species may have become differentiated

by association with different hosts / vectors or geographical isolation. Each speciation event may have had different causes and it is difficult to find a general trend. *Leishmania major* seems associated with arid climates, whilst *L. tropica* is present in semi-arid and cool-arid climates, both found in several countries of Africa and Asia, but some species are more restricted (*L. arabica* in Saudi Arabia, *L. turanica* in Uzbekistan, *L. gerbilli* in China / Mongolia / former USSR and *L. aethiopica* in Ethiopia) (Dereure, 1999). Certain species are host specific (*L. turanica* and *L. gerbilli* in *Rhombomys opimus*), others have preferential hosts (*L. aethiopica* with Procaviidae) and yet others can have several hosts (*L. major*) or are only known in humans and dogs (*L. tropica*) (Dereure, 1999). There is some association of species and vector subgenus: *L. aethiopica* with *P.* (*Larroussius*), *L. major* with *P.* (*Phlebotomus*), *L. tropica* with *P.* (*Paraphlebotomus*) (Esseghir *et al.*, 2000), but not all species in each sub-genus will transmit *Leishmania* and the association does not seem to be of a co-evolutionary nature.

8.4.8. Speculations on the origin of the sub-genus *Leishmania*

There are several competing hypothesis for the origin of *Leishmania (Euleishmania) / Endotrypanum (Paraleishmania)*: Neartic, Paleartic and African origins (see Introduction). *Euleishmania* could have evolved in South America, simultaneously with *Paraleishmania*, and with separation of one branch, *Leishmania* (*Leishmania*), to North America followed by quick spread to the OW, where the *Sauroleishmania* would have evolved (Noyes, 1998b). Alternatively, *Leishmania* (*Leishmania*) could be the most ancient and would have evolved in the OW together with *Sauroleishmania*, and then migrated to the American continent, with *Leishmania* (*Viannia)* and *Paraleishmania* developing in South America (Kerr, 2000). The African hypothesis states that *Leishmania* would already have been present in Southern supercontinent Gondwana, and that the separation of sub-genus *Viannia* and *Leishmania* would have followed separation of Africa and South America (Momen *et al.*, 2000). Section Paraleishmania would have originated after introduction of porcupines in South America.

The Paleartic hypothesis does not explain why *Leishmania* (*Viannia)* and *Paraleishmania* infect mainly mammals from ancient orders in South America and how all phylogenies of *Leishmania / Endotrypanum* are rooted between *Euleishmania* and *Endotrypanum* clades. Neither Paleartic nor African theories explain why *Leishmania* (*Leishmania)* are strongly associated with rodents, which appeared much later than *Paraleishmania* and *Leishmania* (*Viannia)* hosts in the fossil record and

later than the separation of Africa from South America. The African model does not explain why species of *Viannia* are found mainly in tropical forests, whilst species of the sub-genus *Leishmania* are found in mainly savanna like habitats. If indeed *Leishmania* are as old in Africa as they are in South America, it would seem likely that leishmaniasis would be common in African forest environments, as they are in South America. If any ancient *Leishmania* have ever been present in Africa, they must be extinct or be present in an unknown host. Present day medically important parasites must have been imported, forthe reasons stated below. The Neotropical hypothesis thus seems more likely, with *Paraleishmania* and *Leishmania* (*Viannia)* as the first to differentiate in *Leishmania.*

***Euleishmania / Paraleishmania* clade**

From the above considerations and assuming that most phylogenies of *Leishmania* are essentially correct, I believe that the most likely scenario (Fig. 55) is a Neotropical origin for *Leishmania*, but with an early migration of ancestors of the su-genus *Leishmania* and *Sauroleishmania* to North America. Emergence of the clade *Euleishmania / Paraleishmania* must have occurred in South America some time after separation from Africa from a kinetoplastid ancestor, possibly a *Crithidia* like organism, which are monogenetic parasites of Diptera and Hemiptera. The first Phlebotominae, one of the first dipteran groups to individualize, probably appeared in the Middle Jurassic (Lewis, 1982) and it is possible that they were able to spread throughout the supercontinent Pangea before the separation of the two major Northern (Laurasia) and Southern (Gondwana) continents. Later (90 Mya, Late Cretaceous), South America separated from Africa, thus becoming two completely isolated continents (Cox and Moore, 2000). After the separation, a dipteran parasite, probably infecting early Phlebotominae, might have become associated with edentates or marsupials developing a digenetic life cycle. During the early times of Africa-South America separation, caviomorph rodents and some monkeys may have been able to cross to South America (Cox and Moore, 2000). Some *Leishmania* ancestors would have adapted to caviomorphs and differentiated into the *Paraleishmania* clade (Cupolillo *et al.*, 2000).

Cross to North America

During the Earliest Cenozoic (Paleocene, 65 to 56 Mya) 10 different groups of South American mammals were able to migrate to North America through a short lived land bridge (Cox and Moore, 2000). Some of these mammals may even have reached Europe, where the fossil of an anteater (edentate) has been found, and marsupials were able to reach Africa in the Oligocene. A population of *Euleishmania* could have thus migrated from South to North America, probably with edentates,

where they would have first encountered myomorph rodents and would have found other groups of Phlebotominae. Adaptation of some *Leishmania* strains to rodents (and to rodent associated sandflies) might have caused the differentiation between sub-genera *L.* (*Viannia*) and *L.* (*Leishmania*).

The other possible route would be to cross to Africa. There are, however, several objections to this hypothesis. Only the appearance of monkeys and hystricomorphs in South America suggests that there was ever a connection between South America and Africa, which would have occurred no later than the Late Cretaceous or even Early Eocene (Cox and Moore, 2000). There is no evidence, however, of any South American group reaching Africa through this route. The direction of ocean currents is thought to have been from Africa to South America, thus facilitating migration of animals from island to island in that direction but not in the opposite, as a complete land bridge is not thought to have occurred. Furthermore, most *Leishmania* species are present only in the northeast part of Africa and mostly Asia. A btidge into Africa would probably have included tropical forest types as present in modern Central Africa and *Leishmania* would be expected to be present in its whole extension, as it is in South America. A fourth objection will be discussed later.

During the Paleocene / Eocene (65-35Mya), *Leishmania* ancestors, or even some *L.* (*Viannia*) strains, would have been able to disperse in the northern hemisphere and later populations would have been isolated. At that time there were extensive land bridges between North America and both Asia and Europe (Cox and Moore, 2000). Migration of both hosts and vectors would be facilitated by warmer climates (Paleocene with temperature gradients of 5°C from pole to equator) which lasted until the Miocene (23-18Mya). From the beginning of the Miocene the climate became too cool to allow passage through high latitudes, except for larger mammals, and there was complete separation of the three northern continents (Cox and Moore, 2000).

The fourth objection to a crossing to Africa option is that it would not have been possible for an African derived sub-genus *Leishmania* to cross from Asia to North America. Until the Late Oligocene / Early Miocene, Africa was separated from Eurasia (Cox and Moore, 2000). That would mean that *Leishmania* would have to spread to Eurasia and then to North America through the Bering landbridge, which would probably be already too cold for an easy crossing (not many animals crossed, and mostly large ones).

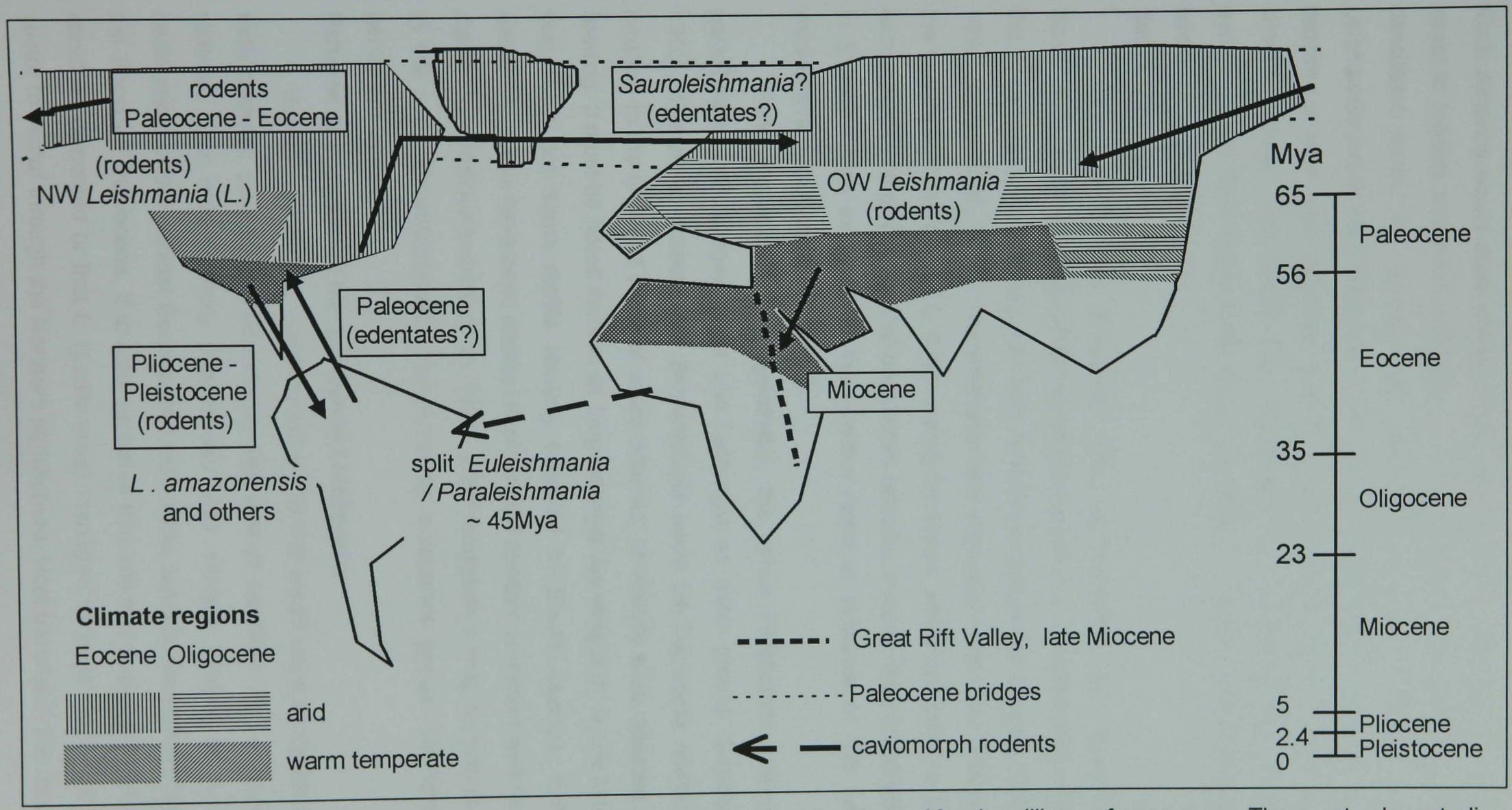


Figure 55 - Illustrated model for evolution of *Leishmania*. See text for explanations. Mya is millions of years ago. The most relevant climate regions are shown (www.scotese.com). Tropical and cool temperate are those nearest to the equator or to the North Pole, respectively, by default.

Here it is proposed that an early separation of the *Leishmania* sub-genus in North America would allow plenty of time and climatic conditions for *Leishmania* to adapt to rodents and cross to the OW. At this point it must be referred that the most consistent model for evolution of *Leishmania* (Noyes, 1998b) proposes that *Leishmania* would have migrated to the OW in the Miocene. However, it would have required a very rapid passage through the Bering land cross during brief milder climate periods. This theory also implies that *L.* (*Sauroleishmania*) would have evolved from OW *Leishmania* which crossed to reptiles and underwent rapid evolutionary change.

Sauroleishmania

The model here proposed also contemplates the hypothesis that *Sauroleishmania* may have indeed separated early from *L.* (*Leishmania*) or even from the same '*Viannia*' ancestors, as some molecular phylogenies suggest (Croan *et al.*, 1997; Noyes *et al.*, 1997). The edentates and marsupials that were able to cross to the OW might have carried *L.* (*Leishmania*) ancestors, which had become associated with rodents in the NW, but that could have become, instead, associated with reptiles in Europe. Upon extinction of the primitive hosts all evidence of NW *Leishmania* would have disappeared in the OW.

Further supporting this hypothesis, the genus *Phlebotomus* seems to be paraphyletic, with *Sergentomyia* and *Lutzomyia* as sister groups (Depaquit *et al.*, 1998). Therefore, present day *Sergentomyia* could be *Lutzomyia* relatives, which crossed to the OW and some of which adapted, or already were adapted, to reptile feeding. It must be noted that not all *Sergentomyia* are obligatory reptile feeders, and that there are some reptile feeding sandflies in South America. Furthermore, *Sauroleishmania* have been isolated from several species of reptiles and have a wide distribution (Killick-Kendrick *et al.*, 1986), which suggests a long evolutionary history. In this case, *Sauroleishmania* should not be a separate genus, but rather a sub-genus of *Leishmania*.

Patterns of evolution in the sub-genus *Leishmania*

Evolution of early *L.* (*Leishmania*) in the NW would have accompanied that of early rodents, probably simultaneously with that of sandflies with a preference for rodents. The rodent family Paramyidae was already spread in the northern hemisphere in the earliest Eocene and most other families would have appeared in the Eocene / Paleocene. It is quite probable that infected populations of rodents or sandflies dispersed or that *L.* (*Leishmania*) managed to 'jump' from host to host or vector to vector through the Northern hemisphere. Host transfer is the most likely, as

Leishmania are less host than vector specific. Separate migrations through either the Bering route or the Greenland route would have contributed to speciation, because Europe and Asia were separated by a shallow sea at the time. Diversification of rodents during the Miocene and ecosystem fragmentation due to aridification could have been a further speciation factor for *Leishmania* (*Leishmania*), especially around the Mediterranean.

With the climate cooling during the Oligocene and Miocene, savanna like habitats, which are the main present day habitat of most hosts of sub-genus *Leishmania* would have become wider. During that time, Africa joined Eurasia which formed again a supercontinent and, in the Miocene, Africa permanently joined Eurasia (Cox and Moore, 2000). Populations of sandflies and rodents would have been pushed south to warmer climates to more or less the present distribution of OW and North American *Leishmania*. Ancient *Leishmania* distributions would have disappeared, thus making it more difficult to determine with exactitude the phylogeny of the sub-genus, which cannot be deduced from fossil records, by association with vector or host phylogenies.

Current distribution and species of OW *Leishmania* are likely to reflect climatic changes of the last 20My, as well as host and vector history. During glacial and inter-glacial periods, populations of parasites would have become isolated in climatic islands or expand, thus producing complicated evolutionary patterns. Speciation events of OW *Leishmania* seem to have occured in both Asia and Africa and some species have evident speciation patterns. The origin of the *L. donovani* complex is one of the most obscure, because different genetic groups have distinct hosts and vector associations. However, and as discussed before, this complex could have differentiated earlier by adaptation to higher temperatures and could have maintained a somewhat low degree of speciation. Some later branches became more host specific, like Indian strains which developed an anthroponotic cycle, and *L. infantum* which speciated with canids. Among CL species two major lines of evolution emerge. One has its hallmark in *L. major.* This species is widely distributed and infects several species of rodents, mainly gerbils, in Asia and Africa and it is considered to be very similar to *L. mexicana*, thus suggesting that *L. major* is the closest relative to *L. mexicana* like OW colonizers. Two related species, *L. turanica* and *L. gerbilli,* may have speciated by geographical isolation from a *L. major* population in gerbils, either from a common ancestor or independently from each other; *L. gerbilli* in China and *L. turanica* in Uzbekistan. Another line of evolution seems to be related to hyraxes (Hyracoidea). *Leishmania killicki,* which is very close to *L. tropica* is associated with Hyracoidea hosts. *Leishmania aethiopica* probably evolved in the Ethiopian region

upon isolation in a life cycle including Hyracoidea (hyraxes) and sandflies of the sub-genus Larroussius. Although closer to *L. major* by IEA typing (Rioux *et al.*, 1990), this species was more related to *L. tropica* in a DNA sequence phylogeny of a repetitive sequence (Piarroux *et al.*, 1995). These uncertain phylogenetic relationships argue in favour of the development of a hyrax evolutionary branch upon colonization of the African continent by *L. major* like *Leishmania*, with an early separation of *L. aethiopica* and *L. killicki. Leishmania tropica* might have evolved from *L. killicki*, the closest species, by adaptation to a human cycle from the hyrax infecting ancestor.

The present day distribution of *Leishmania* species in Africa follows a very clear pattern which also shows that OW *Leishmania* could not have emerged in Africa. Human associated species, such as *L. tropica* and the *L. donovani* complex, but also the generalist species *L. major*, can be found with higher or lesser prevalences, only in savanna like environments of Africa. Until the Miocene (23Mya) Africa was covered by forest (mostly tropical), with the exception of what is now the Sahara desert (www.scotese.com). Only at that time did critecid rodents (strongly associated with *L. major*) invade the continent, but also only then, did shrub and savanna like ecosystems appeare (Cox and Moore, 2000). The formation of the Great Rift Valley, from Central Asia to Mozambique, was accompanied by a rise of the surrounding land, which created a rain shadow to the East. Not only is this major event thought to have potentiated the development of bipedalism in hominids, and thus the emergence of human species, but it also created the arid environments prefered by OW *Leishmania* hosts and vectors and formed a barrier against migration of most species to the western tropical forests. Even today, leishmaniasis is more frequent in countries east of the Rift and only present in regions with similar climates and vegetation, either in Southern Africa or around the Sahara desert, probably due to human action.

A second round of speciations within the *L.* (*Leishmania*) sub-genus might have occurred when North American rodents migrated to South America. Family Sigmodontinae is host to both *L. mexicana* and *L. amazonensis*, and is thought to have invaded South America 9 to 3.5 Mya (Engel, *et al.*, 1998). Other species of the *L. mexicana* complex, such as *L. amazonensis,* might have originated from *L. mexicana* migrants to South America after the establishment of the Panama Isthmus in the Pliocene - Pleistocene, 3-2Mya (Cox and Moore, 2000), and speciations might have occurred, for example, by association with another rodent group, the caviomorphs (such as *L. enrietti* in *Cavia porcellus*), or habitat diversity.

Association of *Leishmania* to humans

It is possible that most OW *Leishmania* species started infecting humans since the appearance of the first hominids. The same savanna regions where hominids were evolving probably were the habitat of the first *Leishmania* to colonize Africa and although most *Leishmania* species are zoonotic, humans can become infected in sylvatic environments. Some groups, such as *L. donovani* (india) and *L. tropica* may have acquired a human cycle, although probably maintaining the capacity to infect other hosts, whilst *L. donovani* (infantile) may have acquired a strong human dependent cycle but on domestic canids. Closer association with humans would be potentiated by sandflies developing anthropophylic feeding habits. African vectors could have followed hominids out of Africa, whilst new associations might have then developed. *Leishmania braziliensis* in the NW could have adapted to a simultaneously human and canid cycle upon the arrival of the first humans in America or even more recently. Speciation of some *Leishmania* in human or canine populations would have been facilitated by their gregarious habits (tribes or villages in humans and packs in wolves / dogs). These human associations, however, pose some phylogenetic problems. The association with human habitats suggests that speciation occurred very recently, within the last hundred of thousands of years for *L. tropica* and antroponotic *L. donovani* and only less than twenty thousands of years for *L. braziliensis*, but both of these species are very diverse, which would suggest a more ancient origin. It would be possible that those species were present in a wild reservoir (probably hyraxes for *L. tropica* and still uncertain rodents or carnivores for *L. donovani*), which disappeared or became secondary upon adaptation to humans. Domestication of the dog probably occurred within the last 20-10 000 years (Wayne, 1993) and so, it is most likely that *L. donovani* (infantile) had adapted to wild canids before domestication of the dog. Until any unknown host is found, or until a reliable molecular clock is found, these questions will remain open.

8.5. Evaluation of the project

The first objective of determining the phylogenetic relationships of *L. chagasi* within the *L. donovani* complex was completely achieved. Using a thorough RAPD analysis of several *L. chagasi*, *L. infantum* and *L. donovani* strains, it was shown here that it is not possible to separate a *L. chagasi* clade. Instead, all tested *L. chagasi* strains belonged to the *L. infantum* clade. Furthermore, the genetic diversity within the *L. infantum* / *L. chagasi* clade was far lower than that within *L. donovani.* Additionally, the partial sequence of the *mspC* gene was identical between the two named species, in contrast to the degree of sequence polymorphism present within *L. donovani*. Analysis of RFLP data fully supported a *L. infantum* / *L. chagasi* clade with low diversity, as did all previous and present isoenzyme analyses. The data pertinent to these objectives and gathered in this project have been published (Mauricio *et al.*, 1999) together with a molecular, epidemiological, clinical and historical defence of synonymy of *L. chagasi* with *L. infantum* (Mauricio *et al.*, 2000).

The second objective, of studying the genetic diversity within *L. infantum* in Portugal was accomplished but not in great detail. Although the Portuguese *L. infantum* strains were found to be polymorphic by RAPD analysis, the resolution power was low and may have prevented detection of clusters corresponding to foci. Indeed no such cluster was observed and if this result is confirmed by future work it means that Portuguese foci are not entirely isolated from each other, which may have important consequences for the control of kala-azar. More sensitive techniques, such as microsatellites may provide more reliable and sensitive markers of population genetics epidemiological studies of such closely related organisms.

The third objective of studying the genetic diversity within the *L. donovani* complex was also achieved with the development of new techniques - *mspC* sequencing and RFLP of ITG/CS and ITG/L - and the adaptation of ITS and mini-exon RFLPs and RAPD. Isoenzyme analysis was used as the reference method. The best typing methods were ITG/CS and ITS RFLPs and *mspC* sequencing, but neither ITG/L nor mini-exon RFLPs produced reliable results.

Based on the different techniques, it was possible to define at least five clusters or clades within the *L. donovani* complex. These groups correlated with IEA typing and some of them also with geographical origin or host specificity. Some clades *L. donovani* were identified that might have arisen through allopatric speciation (india, infantile and red sea), and the contrast between generalist and specialised genetic groups in the history of the complex also became apparent.

The few recently isolated strains present in the analysis indicated that results were valid, despite being based in a majority of older strains. It would however be interesting to broaden this analysis with the inclusion of more recently isolated strains and from more foci, for example from Central Asia, as much genetic diversity may have been missed.

Definition of a sixth clade was tempting but not possible, due to lack of enough strains from China. A future analysis should include not only Chinese strains but also strains from Asian countries previously in the Soviet Union, and more strains from Near and Middle Eastern countries.

There is still much to be resolved on the genetic diversity of the *L. donovani* complex. This project demonstrated that serious genetic diversity studies must include more than one method or technique and also include strains from different origins and phenotypes. Understanding of genetic diversity at the population level is still lacking and would benefit from intensive field studies and the development and application of microsatellite analysis. Importantly, this project has demonstrated that traditional classifications of species in the *L. donovani* complex are wrong. *Leishmania chagasi* is a synonym of *L. infantum* and traditional division of *L. infantum / L. donovani / L. archibaldi* by IEA is not correct. These conclusions will change the design of specific diagnostic methods and the planning of epidemiological studies of the *L. donovani* complex.

A modification for taxonomic nomenclature within the *L. donovani* complex is proposed. Instead of species, to consider the complex as one species, *L. donovani* with the genetic groups defined here named in parentheses: *L. donovani* (india) for *L. donovani s.s.* or from India; *L. donovani* (infantile) for *L. infantum* (*L. chagasi*); *L. donovani* (sudan) for *L. archibaldi* (MON 82), MON 18 and MON 30; *L. donovani* (kenya) for those strains mainly from Kenya; *L. donovani* (red sea) for LON 42 / 52 from coastal areas of the Red Sea; *L. donovani* (china) for MON 35 and other strains which may be found to belong to the same genetic group.

A model for evolution of *Leishmania*, and in particular of Old World *Leishmania* and *Sauroleishmania*, was developed, which can be tested against comprehensive molecular clock phylogenies of sandflies and *Leishmania* and in relation to landmarks such as continent separations and connections and host evolution.

8.6. Concluding remarks

Situation

- Visceral *Leishmania*sis is a life threatening disease estimated to affect 0.5 million people world-wide caused by infection with protozoan parasites of the *L. donovani* complex. The three to four described species (*L. donovani*, *L. infantum*, *L. chagasi* and *L. archibaldi*) are not congruent by isoenzyme typing and epidemiological features and therefore require an extensive revision.

Main achievements

- This thesis clarified the genetic structure of the *L. donovani* complex through application of: newly developed (gp63 intergenic regions) and adapted (ITS and mini-exon) PCR-RFLP; the DNA sequence of a single copy gp63 gene (*mspC*); RAPD and isoenzyme analysis.
- It was here demonstrated that the number of species in the *L. donovani* complex should be greatly reduced. *Leishmania chagasi* and *L. infantum* are synonymous, whilst *L. archibaldi* (zymodeme MON 82) is not a separate species since it did not form a clade of its own. Although the *L. donovani* complex comprised at least five coherent genetic groups (clades), one of which composed of Indian strains and another of *L. infantum* strains, it was considered that these clades should not be considered as species, but as identifiable genetic groups.
- On a smaller scale, it was not possible to identify any kind of genetic structure within Portuguese *L. infantum* strains, suggesting that *L. infantum* may exist in the territory as a single population.

Strains and techniques

- All analyses included a large number of strains belonging to the three main species of the *L. donovani* complex (*L. infantum*, *L. chagasi* and *L. donovani*) or from Portuguese *Leishmania* strains. The chosen strains covered most known or inferred genetic diversity and geographical and host origins.
- Isoenzyme typing was, in this thesis, the validation technique for DNA based typing methods. Isoenzymes generated some genetic diversity, particularly amongst *L. donovani* strains and differentiated between the genetic groups, but were not very useful to produce phylogenies.
- RAPDs provided adequately sensitive data for detection and study of genetic diversity between related strain groups such as *L. infantum* and *L. chagasi* and within homogeneous groups such as Portuguese *L. infantum* strains.

- Sequence of the GP63 coding gene *mspC* provided a sound basis for phylogenetic analysis of the *L. donovani* complex. Two to three main lineages were identified within the complex, one with *L. infantum*, the other with zymodemes LON 42/52 (Group E; Red Sea) and a third with the remainder *L. donovani*; the latter was subdivided in lesser lineages.
- Through both phylogenetic and multivariate analyses of pooled data from PCR-RFLPs of intergenic targets, but also through individual analysis of gp63 ITG/CS PCR-RFLP, the *L. donovani* complex was divided in at least five genetic groups. These five groups were also identified upon analysis of pooled RFLP and RAPD data and correlated with isoenzyme typing and *mspC* DNA sequences.
- Phylogenies of the *L. donovani* complex were difficult to determine, because of the large genetic distance to the outgroups but also because of lack of outgroups in the PCR-RFLPs of gp63 intergenic regions.

Re-evaluation of the *L. donovani* complex

- *Leishmania chagasi* and *L. infantum* are synonymous. Despite the geographical separation, *L. chagasi* is genetically indistinguishable from *L. infantum* and, as discussed in this thesis, its introduction in the New World must have been very recent (within the last 500 years).
- The *Leishmania donovani* complex comprises at least five genetic groups of strains at least as diverse as *L. infantum* and named as genetic groups of *L. donovani*: *Leishmania infantum* (*L. chagasi*), named as *L. donovani* (infantile); Indian strains, as *L. donovani* (india); 'Sudan' strains (MON 18/30/82), as *L. donovani* (sudan); 'Kenyan' strains, as *L. donovani* (kenya); and 'Ethiopian / Saudi Arabian' strains (LON 42/52), as *L. donovani* (red sea). A strain from China (Wangjie 1) might belong to a fifth group, named *L. donovani* (china) but more strains should be analyzed. It was not found appropriate that these groups should be given the species status, but their recognition may provide a better working basis for epidemiological studies of *L. donovani* and development of more effective diagnostic methods.
- Characterization of the genetic groups within the *L. donovani* complex has important implications to several aspects of leishmaniasis control. Specific diagnosis and strain typing, in particular to differentiate *L. infantum*, will benefit from a clear genetic definition of strain groups. For example, strains which have been considered as *L. infantum* until now, must be eliminated from specificity assays and researchers can be aware of those '*L. donovani*' groups that are most likely to be confused with *L. infantum*.

Retrospective

- In retrospect, it would have been interesting to adapt and analyse enzymes used for typing in Montpellier to achieve better correlation between the zymodemes determined here and those described by Rioux *et al.* (1990). Although those extra enzymes would have added more data to determine better the isoenzyme phylogeny of the *L. donovani* complex, the analyses done using the Montpellier isoenzyme data did not seem very robust and thus the effort might not have been worthwhile.
- RAPD analysis might have been improved by increased annealing temperature or use of longer primers, and thus by increased specificity. Band scoring of RAPD profiles would have been more sound and accurate if homologous bands had been determined by probing Southern blots of RAPD profiles with selected RAPD bands.
- Phylogenetic analysis of RFLP data could have been more accurate if variant PCR products had been sequenced to identify restriction sites or the cause of fragment size differences. Use of DGGE, or other methods, would have also been helpful to determine the number of alleles present for each RFLP target. Those further analyses were not done because of the exaggerated extra work and cost.
- Regarding the DNA sequencing of *mspC*, strain Wangjie 1 was not sequenced.

Future work

- Some open chapters on the *L. donovani* complex were closed in this thesis, but some new questions also arose, which will need to be addressed in the future.
- In the future, the taxonomic position of Wangjie 1 will be clarified by analysis of the DNA sequence of the *mspC* gene.
- Potential diagnostic targets emerged from this project, such as *mspC* single nucleotide polymorphisms, certain RAPD bands and RFLPs. It would also be useful to characterize certain *L. infantum* or *L. donovani* specific RAPD bands which could be used as probes or as PCR targets for differential diagnosis or typing of strains within the *L. donovani* complex. Although sequencing is still expensive and time consuming, some single nucleotide polymorphisms were identified within *mspC* which could be typed using fast diagnostic methods such as TaqMan and molecular beacons or could form part of DNA microarrays for typing.
- Having clarified the genetic diversity structure of *L. donovani*, it is necessary to investigate if the genetic groups of *L. donovani*, have biological or epidemiological significance. More specifically, if those genotypes correlate with differential pathogenicity, drug resistance or vector/host associations.

- This project raised the question of whether genetic recombination occurs undetected in the *L. donovani* complex. By using more sensitive typing methods, such as microsatellites, heterozygotes could be detected in natural populations, whilst by using defined markers, such as single nucleotide polymorphisms in *mspC*, hybridization events might be detected by experimental crosses in the laborarory.
- The PCR-RFLP techniques and the DNA sequence of *mspC* (or other GP63 genes) as used here, may contribute towards investigation of the origin and evolution of the sub-genus *Leishmania* in the New and Old Worlds, by application to a large range of strains from all *Leishmania* (*Leishmania*) species and by pooling with or comparing to other methods or genomic targets.

9. References

Abranches, P., Conceição Silva, F. M. and Silva Pereira, M. C. (1984) Kala-azar in Portugal. V. The sylvatic cycle in the enzootic endemic focus of Arrábida. *Journal of Tropical Medicine and Hygiene* **87,** 197-200.

Abranches, P., Pires, C. A., Conceição-Silva, F. M., Silva-Pereira, M. C. D. and Gomes, G. M. S. (1987) O kala-azar em Portugal - VI. Inquérito epidemiológico realizado na região metropolitana de Lisboa: interpretação da estrutura e dinâmica do foco endémico. *Jornal das Ciências Médicas de Lisboa* **151,** 364-379.

Abranches, P., Ribeiro, M. S., Lopes, F. J. and Gomes, L. L. T. (1983) O kala-azar em Portugal. II. Inquérito sobre leishmaniose em Alcácer do Sal. *Jornal da Sociedade de Ciências Médicas de Lisboa,* 147-149.

Abranches, P., Sampaio-Silva, M. L., Santos-Gomes, G. M., Avelino, I. C., Pires, C. A., Conceição-Silva, F. M., Seixas-Lopes, A., Silva-Pereira, M. C. D. and Janz, J. G. (1992) Kala-azar in Portugal. VII. Epidemiological survey in Alijó (Endemic Region of Alto Douro) *Research and Reviews in Parasitology* **52,** 121-124.

Agatsuma, T. (1992) [Molecular biology of *Leishmania* and diagnosis of leishmaniasis by DNA probe]. *Nippon Rinsho* **50,** S469-474.

al Jurayyan, N. A., al Ayed, I. H., al Nasser, M. N., al Mugeiren, M. M., Boohene, A. G. and al Herbish, A. S. (1992) Visceral leishmaniasis in infancy and childhood epidemiology and clinicopathological study of 63 cases in Al-Baha Province, Saudi Arabia. *Journal of Tropical Paediatrics* **38,** 12-16.

Alvar, J., Ballesteros, J. A., Soler, R., Benito, A., van Eys, G. J., Schoone, G. J. and Cabrer, B. (1990) Mucocutaneous leishmaniasis due to *Leishmania* (*Leishmania*) *infantum*: biochemical characterization. *American Journal of Tropical Medicine and Hygiene* **43,** 614-618.

Alvarez, F., Cortinas, M. N. and Musto, H. (1996) The analysis of protein coding genes suggests monophyly of *Trypanosoma*. *Molecular Phylogenetics and Evolution* **5,** 333-343.

Alvarez, F., Robello, C. and Vignali, M. (1994) Evolution of codon usage and base contents in kinetoplastid protozoans. *Molecular Biology and Evolution* **11,** 790-802.

Alves Pires, C. and Ribeiro, H. (1991) The phlebotomine sandflies of Portugal. V. Observations on the ecology of the vectors of leishmaniasis in the Alto Douro region. *Parassitologia* **33,** 63-68.

Andresen, K., Ibrahim, M. E., Theander, T. G. and Kharazmi, A. (1996) Random amplified polymorphic DNA for the differentiation of *Leishmania donovani* isolates from Sudan. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **90,** 204-205.

Antoine, J.-C., Lang, T. and Prina, E. (1999) Biologie cellulaire de *Leishmania*. *In Les leishmanioses (ed.* Dedet, J. P.), Ellipses, pp. 41-58.

Antoine, J.-C., Prina, E., Lang, T. and Courret, N. (1998) The biogenesis and properties of the parasitophorous vacuoles that harbour *Leishmania* in murine macrophages.

Trends in Microbiology **6,** 392-341.

Ashford, R. W. and Bates, P. A. (1998) Leishmaniasis in the Old World. In *Parasitology* (Eds. Cox, F. E. G., Kreier, J. P. and Wakelin, D.), Arnold, pp. 215-240.

Avise, J. C. (1994) *Molecular markers, natural history and evolution*. Chapman & Hall, New York. 511 pp.

Bansal, S. K. and Singh, K. V. (1996) Susceptibility status of *Phlebotomus papatasi* and *Sergentomyia punjabaensis* (Diptera: Psychodidae) to some insecticides in district Bikaner (Rajasthan) *Journal of Communicable Diseases* **28,** 28-32.

Banuls, A. L., Brisse, S., Sidibe, I., Noel, S. and Tibayrenc, M. (1999a) A phylogenetic analysis by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting of the *Leishmania* genome project Friedlin reference strain. *Folia Parasitologica (Praha)* **46,** 10-14.

Banuls, A. L., Hide, M. and Tibayrenc, M. (1999b) Molecular epidemiology and evolutionary genetics of *Leishmania* parasites. *International Journal for Parasitology* **29,** 1137-1147.

Bao, Y., Wang, S. T. and Shao, Q. F. (1994) A further study of LDT and IFAT tests in evaluating the control of kala-azar in China. *Journal of Tropical Medicine and Hygiene* **97,** 357-361.

Barrios, M., Rodriguez, N., Feliciangeli, D. M., Ulrich, M., Telles, S., Pinardi, M. E. and Convit, J. (1994) Coexistence of two species of *Leishmania* in the digestive tract of the vector *Lutzomyia ovallesi*. *American Journal of Tropical Medicine and Hygiene* **51,** 669-675.

Bastien, P., Blaineau, C. and Pages, M. (1992) *Leishmania*: sex, lies and karyotype. *Parasitology Today* **8,** 174-176.

Bates, P. A. (1994) Complete developmental cycle of *Leishmania mexicana* in axenic culture. *Parasitology* **108,** 1-9.

Belkaid, Y., Kamhawi, S., Modi, G., Valenzuela, J., Noben Trauth, N., Rowton, E., Ribeiro, J. and Sacks, D. L. (1998) Development of a natural model of cutaneous leishmaniasis: powerful effects of vector saliva and saliva preexposure on the long-term outcome of *Leishmania major* infection in the mouse ear dermis. *Journal of Experimental Medicine* **188,** 1941-1953.

Belli, A., Rodriguez, B., Aviles, H. and Harris, E. (1998) Simplified polymerase chain reaction detection of new world *Leishmania* in clinical specimens of cutaneous leishmaniasis. *American Journal of Tropical Medicine and Hygiene* **58,** 102-109.

Belli, A. A., Miles, M. A. and Kelly, J. M. (1994) A putative *Leishmania panamensis / Leishmania braziliensis* hybrid is a causative agent of human cutaneous leishmaniasis in Nicaragua. *Parasitology* **109,** 435-442.

Belova, E. M. (1971) Reptiles and their importance in the epidemiology of leishmaniasis. *Bulletin of the World Health Organization* **44,** 553-560.

Ben Ismail, R., Smith, D. F., Ready, P. D., Ayadi, A., Gramiccia, M., Ben Osman, A. and Ben Rachid, M. S. (1992) Sporadic cutaneous leishmaniasis in north Tunisia:

identification of the causative agent as *Leishmania infantum* by the use of a diagnostic deoxyribonucleic acid probe. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86,** 508-510.

Benavides, G. R., Sullivan, J. J., Steurer, F., McGraw, R. A. and Tarleton, R. L. (1993) Differentiation of trypanosomatid species by hybridization to selected rRNA probes. *Molecular and Cellular Probes* **7,** 89-96.

Berman, J. D. (1997) Human leishmaniasis: clinical, diagnostic, and chemotherapeutic developments in the last 10 years. *Clinical Infectious Diseases* **24,** 684-703.

Berman, J. D. and Neva, F. A. (1981) Effect of temperature on multiplication of *Leishmania* amastigotes within human monocyte-derived macrophages in vitro. *American Journal of Tropical Medicine and Hygiene* **30,** 318-321.

Beverley, S. M., Ismach, R. and McMahon-Pratt, D. (1984) Evolution of the genus *Leishmania* Ross, 1903 as revealed by comparisons of nuclear DNA restriction fragment patterns. In *Leishmania. Taxonomie et phylogenese. Applications eco-epidemiologiques.*(Ed. Rioux, J.P.) *Montpellier*, pp. 265-267.

Beverley, S. M., Ismach, R. B. and Pratt, D. M. (1987) Evolution of the genus *Leishmania* as revealed by comparisons of nuclear DNA restriction fragment patterns. *Proceedings of the National Academy of Sciences USA* **84,** 484-488.

Bhattacharyya, R., Singh, R., Hazra, T. K. and Majumder, H. K. (1993) Application of polymerase chain reaction with specific and arbitrary primers to identification and differentiation of *Leishmania* parasites. *FEMS Microbiology Letters* **114,** 99-104.

Binhazim, A. A., Githure, J. I., Muchemi, G. K. and Reid, G. D. (1987) Isolation of *Leishmania major* from a naturally infected vervet monkey (*Cercopithecus aethiops*) caught in Kiambu District, Kenya. *Journal of Parasitology* **73,** 1278-1279.

Blackwell, J. M. (1992) Leishmaniasis epidemiology: all down to the DNA. *Parasitology* **104,** S19-34.

Blum, J., Junghanss, T. and Hatz, C. (1994) Irrwege in der Diagnostik kutaner und mukokutaner Leishmaniosen. *Schweizerische Rundschau Fur Medizin Praxis* **83,** 1025-9.

Bray, R. S. and Lainson, R. (1966) The immunology and serology of leishmanisis. IV. Results of Ouchterlony double diffusion tests. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **60,** 605-609.

Breniere, S. F., Telleria, J., Bosseno, M. F., Buitrago, R., Bastrenta, B., Cuny, G., Banuls, A. L., Brewster, S. and Barker, D. C. (1999) Polymerase chain reaction-based identification of New World *Leishmania* species complexes by specific kDNA probes. *Acta Tropica* **73,** 283-293.

Briones, M. R., Nelson, K., Beverley, S. M., Affonso, H. T., Camargo, E. P. and Floeter Winter, L. M. (1992) *Leishmania tarentolae* taxonomic relatedness inferred from phylogenetic analysis of the small subunit ribosomal RNA gene. *Molecular and Biochemical Parasitology* **53,** 121-127.

Brittingham, A., Morrison, C. J., McMaster, W. R., McGwire, B. S., Chang, K. P. and Mosser, D. M. (1995) Role of the *Leishmania* surface protease gp63 in complement fixation, cell adhesion, and resistance to complement-mediated lysis. *Journal of Immunology* **155,** 3102-3111.

Bryceson, A. D. M. (1996) Chapter 65. Leishmaniasis. In *Manson's Tropical Diseases* (Ed. Cook, G. C.), pp. 1213-1245.

Button, L. L., Russell, D. G., Klein, H. L., Medina Acosta, E., Karess, R. E. and McMaster, W. R. (1989) Genes encoding the major surface glycoprotein in *Leishmania* are tandemly linked at a single chromosomal locus and are constitutively transcribed. *Molecular and Biochemical Parasitology* **32,** 271-283.

Cabral, M., JE, O. G., Gomes, S., Sousa, J. C., Thompson, H. and Alexander, J. (1998) The immunology of canine leishmaniosis: strong evidence for a developing disease spectrum from asymptomatic dogs. *Veterinary Parasitology* **76,** 173-180.

Callahan, H. L., Portal, I. F., Bensinger, S. J. and Grogl, M. (1996) *Leishmania* spp: temperature sensitivity of promastigotes in vitro as a model for tropism in vivo. *Experimental Parasitology* **84,** 400-9.

Camin, J.H. and Sokal, R.R. (1965) A method for deducing branching sequences in phylogeny. *Evolution.* **19**, 311-326

Campbell, D. A., Kubo, K., Clark, C. G. and Boothroyd, J. C. (1987) Precise identification of cleavage sites involved in the unusual processing of trypanosome ribosomal RNA. *Journal of Molecular Biology* **196,** 113-124.

Campino, L., Rica-Capela, M. J., Mauricio, I. L., Ozensoy, S. and Abranches, P. (1995) O kala-azar em Portugal. IX. A região do Algarve: Inquérito epidemiológico sobre o reservatório canino no concelho de Loulé. *Revista Portuguesa de Doenças Infecciosas* **18,** 189-194.

Campino, L., Santos Gomes, G., Pratlong, F., Dedet, J. P. and Abranches, P. (1994) The isolation of *Leishmania donovani* MON-18, from an AIDS patient in Portugal: possible needle transmission. *Parasite* **1,** 391-392.

Campino, L., Santos Gomes, G. M., Pratlong, F., Antunes, F., Mauricio, I., Dedet, J. P. and Abranches, P. (1997) HIV/*Leishmania* co-infections in Portugal: diagnosis and isoenzyme characterization of *Leishmania*. *Annals of Tropical Medicine and Parasitology* **91,** 433-436.

Cariello, N. F., Scott, J. K., Kat, A. G., Thilly, W. G. and Keohavong, P. (1988) Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: HPRT Munich. *American Journal of Human Genetics* **42,** 726-734.

Carrera, L., Gazzinelli, R. T., Badolato, R., Hieny, S., Muller, W., Kuhn, R. and Sacks, D. L. (1996) *Leishmania* promastigotes selectively inhibit interleukin 12 induction in bone marrow-derived macrophages from susceptible and resistant mice. *Journal of Experimental Medicine* **183,** 515-526.

Cerf, B. J., Jones, T. C., Badaro, R., Sampaio, D., Teixeira, R. and Johnson, W. D., Jr. (1987) Malnutrition as a risk factor for severe visceral leishmaniasis. *Journal of Infectious Diseases* **156,** 1030-1033.

Chakrabarty, R., Mukherjee, S., Lu, H. G., McGwire, B. S., Chang, K. P. and Basu, M. K. (1996) Kinetics of entry of virulent and avirulent strains of *Leishmania donovani* into macrophages: a possible role of virulence molecules (gp63 and LPG) *Journal of Parasitology* **82,** 632-635.

Chakrabarti, A., Kumar, B., Das, A. and Mahajan, V. K. (1997) Atypical post kala-azar dermal leishmaniasis resembling histoid leprosy. *Leprosy Review* **68,** 247-251.

Chance, M. L. (1995) New developments in the chemotherapy of leishmaniasis. *Annals Tropical Medicine and Parasitology* **89** S1, 37-43.

Chavez, J. and Witremundo Torrealba, J. (1965) Leishmaniasis visceral humana: su diagnostico por la reaccion quantitative de fijacion del complemento en gotas. *Acta Cientifica Venezoelana* **16,** 213-214.

Chetverin, A. B. and Kramer, F. R. (1993) Sequencing of pools of nucleic acids on oligonucleotide arrays. *Biosystems* **30,** 215-231.

Cillari, E., Vitale, G., Arcoleo, F., P, D. A., Mocciaro, C., Gambino, G., Malta, R., Stassi, G., Giordano, C., Milano, S. *et al.* (1995) In vivo and in vitro cytokine profiles and mononuclear cell subsets in Sicilian patients with active visceral leishmaniasis. *Cytokine* **7,** 740-745.

Comeau, A. M., Miller, S. I. and Wirth, D. F. (1986) Chromosome location of four genes in *Leishmania*. *Molecular and Biochemical Parasitology* **21,** 161-169.

Corliss, J.O. (1998) Chapter 22. Classification of protozoa and protists: the current status. In *Evolutionary Relationships Among Protozoa* (eds. Coombs, G. H., Vickerman, K., Sleigh, M. A., Warren, A.). Kluwer Academic Publishers, Dordrecht, pp 409-447

Cox, C.B. and Moore, P.D. (2000) *Biogeography. An ecological and evolutionary approach*. 6th edition. Blackwell Science, London. 298 pp

Croan, D. G., Morrison, D. A. and Ellis, J. T. (1997) Evolution of the genus *Leishmania* revealed by comparison of DNA and RNA polymerase gene sequences. *Molecular and Biochemical Parasitology* **89,** 149-159.

Cruz, A., Coburn, C. M. and Beverley, S. M. (1991) Double targeted gene replacement for creating null mutants. *Proceedings of the National Academy of Sciences USA* **88,** 7170-7174.

Cuba, C. A., Torno, C. O., Ledesma, O., Visciarelli, E., Garcia, S., Prat, M. I., Costamagna, R., Barbieri, L. and Evans, D. A. (1996) Human cutaneous leishmaniasis caused by *Leishmania* (*Viannia*) *braziliensis* in Santiago del Estero, Argentina: identification of parasites by monoclonal antibodies and isoenzymes. *Revista do Instituto de Medicina Tropical de São Paulo* **38,** 413-421.

Cunha, A. M. (1938) Infecções experimentaes na Leishmaniose visceral americana *Memórias do Instituto Oswaldo Cruz* **33,** 581-598.

Cupolillo, E., Grimaldi, G., Jr. and Momen, H. (1994) A general classification of New World *Leishmania* using numerical zymotaxonomy. *American Journal of Tropical Medicine and Hygiene* **50,** 296-311.

Cupolillo, E., Grimaldi Junior, G., Momen, H. and Beverley, S. M. (1995) Intergenic region typing (IRT): a rapid molecular approach to the characterization and evolution of *Leishmania*. *Molecular and Biochemical Parasitology* **73,** 145-155.

Cupolillo, E., Medina-Acosta, E., Noyes, H., Momen, H. and Grimaldi, G. J. (2000) A revised classification for *Leishmania* and *Endotrypanum*. *Parasitology Today* **16,** 142-144.

Cupolillo, E., Momen, H. and Grimaldi, G., Jr. (1998) Genetic diversity in natural populations of New World *Leishmania*. *Memórias do Instituto Oswaldo Cruz* **93,** 663-668.

Davidson, R. N. (1998) Practical guide for the treatment of leishmaniasis. *Drugs* **56,** 1009-1018.

De Beer, P., El Harith, A., Deng, L. L., Semiao Santos, S. J., Chantal, B. and Van Grootheest, M. (1991) A killing disease epidemic among displaced Sudanese population identified as visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* **44,** 283-289.

De Lange, T., Liu, A. Y., Van der Ploeg, L. H., Borst, P., Tromp, M. C. and Van Boom, J. H. (1983) Tandem repetition of the 5' mini-exon of variant surface glycoprotein genes: a multiple promoter for VSG gene transcription? *Cell* **34,** 891-900.

De Souza Leão, S., Lang, T., Prina, E., Hellio, R. and Antoine, J. C. (1995) Intracellular *Leishmania amazonensis* amastigotes internalise and degrade MHC class II molecules of their host cells. *Journal of Cell Science* **108,** 3219-3231.

Decker-Jackson, J. E., Schrot, J. R. and Levin, G. V. (1977) Identification of *Leishmania spp.* by radiorespirometry. *Journal of Protozoology* **24,** 463-470.

Decker-Jackson, J. E. and Tang, D. B. (1982) Identification of *Leishmania spp.* by radiorespirometry II. A statistical method of data analysis to evaluate the reproducibility and sensitivity of the technique. In *Biochemical characterization of Leishmania.* Proceedings of a Workshop held at the Pan American Health Organization, 9-11 December 1980. UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, Geneva. pp. 205-245.

Degrave, W., Fernandes, O., Campbell, D., Bozza, M. and Lopes, U. (1994) Use of molecular probes and PCR for detection and typing of *Leishmania* - a mini-review. *Memórias do Instituto Oswaldo Cruz* **89,** 463-469.

Del Giudice, P., Marty, P., Lacour, J. P., Perrin, C., Pratlong, F., Haas, H., Dellamonica, P. and Le Fichoux, Y. (1998) Cutaneous leishmaniasis due to *Leishmania infantum*. Case reports and literature review. *Archives of Dermatology* **134,** 193-198.

Delgado, O., Cupolillo, E., Bonfante Garrido, R., Silva, S., Belfort, E., Grimaldi Junior, G. and Momen, H. (1997) Cutaneous leishmaniasis in Venezuela caused by infection

with a new hybrid between *Leishmania* (*Viannia*) *braziliensis* and *L.* (*V.*) *guyanensis*. *Memórias do Instituto Oswaldo Cruz* **92,** 581-582.

Depaquit, J., Perrotey, S., Lecointre, G., Tillier, A., Tillier, S., Ferte, H., Kaltenbach, M. and Leger (1998) Systematique moleculaire des Phlebotominae: etude pilote. Paraphylie du genre Phlebotomus. *Comptes Rendus de L'Academie de Sciences de Paris. Sciences de la vie*. **321**, 849-855

Dereure, J. (1999) Reservoirs de leishmanies. In *Les leishmanioses*, (Ed. Dedet, J. P.) Ellipses, pp. 107-130.

Desjeux, P. (1992) Human leishmaniases: epidemiology and public health aspects. *World Health Statistics Quaterly* **45,** 267-275.

Desowitz, R. S., Draper, C. C. and Phillips, T. M. (1975) Proceedings: Preliminary observations on countercurrent immuno-electrophoresis of *Trypanosoma cruzi* and *Leishmania*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **69,** 430.

Dhar, S., Kaur, I., Dawn, G., Sehgal, S. and Kumar, B. (1995) Post kala-azar dermal leishmaniasis mimicking leprosy: experience with 4 patients, with some unusual features in 1. *Leprosy Review* **66,** 250-256.

Docampo, R. and Moreno, S. N. J. (1999) Acidocalcisome: a novel Ca^{2+} storage compartment in trypanosomatids and apicomplexan parasites. *Parasitology Today* **15,** 443-448.

Docampo, R., Scott, D. A., Vercesi, A. E. and Moreno, S. N. (1995) Intracellular Ca^{2+} storage in acidocalcisomes of *Trypanosoma cruzi*. *Biochemical Journal* **310,** 1005-1012.

Donnelly, K. B., Lima, H. C. and Titus, R. G. (1998) Histologic characterization of experimental cutaneous leishmaniasis in mice infected with *Leishmania braziliensis* in the presence or absence of sand fly vector salivary gland lysate. *Journal of Parasitology* **84,** 97-103.

Dos Santos, J. I., Morgado, M. G. and Galvao Castro, B. (1987) Human visceral leishmaniasis: analysis of the specificity of humoral immune response to polypeptides of *Leishmania donovani chagasi*. *American Journal of Tropical Medicine and Hygiene* **37,** 263-270.

Dujardin, J. C., Banuls, A. L., Llanos Cuentas, A., Alvarez, E., DeDoncker, S., Jacquet, D., Le Ray, D., Arevalo, J. and Tibayrenc, M. (1995a) Putative *Leishmania* hybrids in the Eastern Andean valley of Huanuco, Peru. *Acta Tropica* **59,** 293-307.

Dujardin, J. C., Banuls, A. L., Victoir, K., DeDoncker, S., Arevalo, J., Llanos Cuentas, A., Tibayrenc, M. and Le Ray, D. (1995b) From population to genome: ecogenetics of *Leishmania* (*Viannia*) *braziliensis* and *L. (V.) peruviana*. *Annals of Tropical Medicine and Parasitology* **89** S1, 45-53.

Dujardin, J. C., Dujardin, J. P., Tibayrenc, M., Timperman, G., De Doncker, S., Jacquet, D., Arevalo, J., Llanos Cuentas, A., Guerra, H., Bermudez, H., Hamers, R. and Le Ray, D.

(1995c) Karyotype plasticity in neotropical *Leishmania*: an index for measuring genomic distance among *L. (V.) peruviana* and *L. (V.) braziliensis* populations. *Parasitology* **110,** 21-30.

Dunan, S., Frommel, D., Monjour, L., Ogunkolade, B. W., Cruz, A. and Quilici, M. (1989) Vaccination trial against canine visceral leishmaniasis. Phocean Veterinary Study Group on Visceral Leishmaniasis. *Parasite Immunology* **11,** 397-402.

Dursunova, S. M., Karapetyan, A. B. and Ponirovsky, E. N. (1965) Nekotorye dannye po izycheniiu vistseral'nogo leishmanioza v turkmenskoi SSR. *Medskaya Parazit Bolezni* **4,** 303.

Dye, C. and Williams, B. G. (1993) Malnutrition, age and the risk of parasitic disease: visceral leishmaniasis revisited. *Proceedings of the Royal Society of London B Biological Sciences* **254,** 33-39.

Eisenberger, C. L. and Jaffe, C. L. (1999) *Leishmania*: identification of Old World species using a permissively primed intergenic polymorphic-polymerase chain reaction. *Experimental Parasitology* **91,** 70-77.

El Azzouni, M. Z., Negm, A. Y., El Naga, I. F. and El Din, D. M. (1998) Study and characterization of different developmental stages of *Leishmania major*. *Journal of the Egyptian Society of Parasitology* **28,** 753-763.

El Harith, A., Slappendel, R. J., Reiter, I., van Knapen, F., de Korte, P., Huigen, E. and Kolk, A. H. (1989) Application of a direct agglutination test for detection of specific anti-*Leishmania* antibodies in the canine reservoir. *Journal of Clinical Microbiology* **27,** 2252-2257.

El Hassan, A. M., Meredith, S. E., Yagi, H. I., Khalil, E. A., Ghalib, H. W., Abbas, K., Zijlstra, E. E., Kroon, C. C., Schoone, G. J. and Ismail, A. (1995) Sudanese mucosal leishmaniasis: epidemiology, clinical features, diagnosis, immune responses and treatment. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **89,** 647-652.

Elias, M., Rahman, A. J. and Khan, N. I. (1989) Visceral leishmaniasis and its control in Bangladesh. *Bulletin of the World Health Organization* **67,** 43-49.

Ellis, J. T. and Crampton, J. M. (1991) Differences between *Leishmania (Leishmania) chagasi*, *L. (L.) infantum* and *L. (L.) donovani* as shown by DNA fingerprinting. *Memórias do Instituto Oswaldo Cruz* **86,** 479-481.

Engel, S., Hogan, K., Taylor, J. and Davis, S. (1998) Molecular systematics and paleobiogeography of the South American sigmodontine rodents. *Molecular Biology and Evolution* **15,** 35-49.

Esseghir, S., Ftaiti, A., Ready, P. D., Khadraoui, B., Zaafouri, B., Dellagi, K. and Ben Ismail, R. (1993) The squash blot technique and the detection of *Leishmania major* in *Phlebotomus papatasi* in Tunisia. *Archives de l'Institut Pasteur de Tunis* **70,** 493-496.

Esseghir, S., Ready, P.D. and Ben-Ismail, R. (2000) Speciation of *Phlebotomus* sandflies of the sub-genus *Larroussius* coincided with the late Miocene-Pliocene aridification of

the Mediterranean region. *Biological Journal of the Linnean Society* **70**, 189-219.

Evans, D., Godfrey, D., Lanham, S., Lanotte, G., Modabber, F. and Schnur, L. (1989) *Handbook on isolation, characterization and cryopreservation of Leishmania.* WHO.

Evans, D. A., Kennedy, W. P., Elbihari, S., Chapman, C. J., Smith, V. and Peters, W. (1987) Hybrid formation within the genus *Leishmania*? *Parassitologia* **29,** 165-173.

Evans, D. A., Lanham, S. M., Baldwin, C. I. and Peters, W. (1984) The isolation and isoenzyme characterization of *Leishmania braziliensis* subsp. from patients with cutaneous leishmaniasis acquired in Belize. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **78,** 35-42.

Evans, T. G., Reed, S. S. and Hibbs, J. B., Jr. (1996) Nitric oxide production in murine leishmaniasis: correlation of progressive infection with increasing systemic synthesis of nitric oxide. *American Journal of Tropical Medicine and Hygiene* **54,** 486-489.

Evans, T. G., Teixeira, M. J., Sousa, A. d. Q. and Pearson, R. D. (1995) Short report: extended follow-up of the natural history of persons infected with *Leishmania chagasi. American Journal of Tropical Medicine and Hygiene* **53,** 360-361.

Farris, J.S. (1970) Methods for computing Wagner trees. *Systematic Zoology*. **19**, 83-92.

Farris, J.S. (1977) Phylogenetic analysis under Dollo's law. *Systematic Zoology*. **26**, 77-88.

Fehniger, T. E., Mengistu, G., Gessesse, A., Gabre Mariam, H. and Akuffo, H. (1990) Changes in the antigenic profile of *Leishmania* parasites following shifts in temperature. *Acta Tropica* **47:** 227-236.

Felsenstein, J. (1993) PHYLIP (Phylogeny Inference Package), Distributed by the author. Department of Genetics, University of Washington.

Felsenstein, J. (1988) Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics* **22,** 521-565.

Fernandes, O., Murthy, V. K., Kurath, U., Degrave, W. M. and Campbell, D. A. (1994) Mini-exon gene variation in human pathogenic *Leishmania* species. *Molecular and Biochemical Parasitology* **66,** 261-271.

Fitch, W.M. and Margoliash, E. (1967) Construction of phylogenetic trees. *Science*. **155**, 279-284.

Flinn, H. M., Rangarajan, D. and Smith, D. F. (1994) Expression of a hydrophilic surface protein in infective stages of *Leishmania major. Molecular and Biochemical Parasitology* **65,** 259-270.

Gaafar, A., Kharazmi, A., Ismail, A., Kemp, M., Hey, A., Christensen, C. B., Dafalla, M., El Kadaro, A. Y., El Hassan, A. M. and Theander, T. G. (1995) Dichotomy of the T cell response to *Leishmania* antigens in patients suffering from cutaneous leishmaniasis; absence or scarcity of Th1 activity is associated with severe infections. *Clinical and Experimental Immunology* **100,** 239-245.

Galrito, M., Serrano, A., Dinis, F. and Goncalves, H. (1991) Kala-azar infantil. A propósito de três casos clinicos. *O Petiz* **1,** 37-41.

Garcia, A. L., Carrasco, H. J., Schofield, C. J., Stothard, J. R., Frame, I. A., Valente, S. A.

and Miles, M. A. (1998) Random amplification of polymorphic DNA as a tool for taxonomic studies of triatomine bugs (Hemiptera: Reduviidae) *Journal of Medicine and Entomology* **35,** 38-45.

Gardener, P. J., Chance, M. L. and Peters, W. (1974) Biochemical taxonomy of *Leishmania*. II: Electrophoretic variation of malate dehydrogenase. *Annals of Tropical Medicine and Parasitology* **68,** 317-325.

Ghalib, H. W., Eltoum, E. A., Kroon, C. C. and El Hassan, A. M. (1992) Identification of *Leishmania* from mucosal leishmaniasis by recombinant DNA probes. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86,** 158-160.

Ghosh, K. N. and Mukhopadhyay, J. (1998) The effect of anti-sandfly saliva antibodies on *Phlebotomus argentipes* and *Leishmania donovani*. *International Journal for Parasitology* **28,** 275-281.

Giannini, S. H., Schittini, M., Keithly, J. S., Warburton, P. W., Cantor, C. R. and Van der Ploeg, L. H. (1986) Karyotype analysis of *Leishmania* species and its use in classification and clinical diagnosis. *Science* **232,** 762-765.

Godfrey, D. G. and Kilgour, V. (1976) Enzyme electrophoresis in characterising the causative organism of Gambian trypanosomiasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **70,** 219-224.

Gomes, R. F., Macedo, A. M., Pena, S. D. and Melo, M. N. (1995) *Leishmania* (*Viannia*) *braziliensis*: genetic relationships between strains isolated from different areas of Brazil as revealed by DNA fingerprinting and RAPD. *Experimental Parasitology* **80,** 681-687.

Gonzalez Aseguinolaza, G., Almazan, F., Rodriguez, J. F., Marquet, A. and Larraga, V. (1997) Cloning of the gp63 surface protease of *Leishmania infantum*. Differential post-translational modifications correlated with different infective forms. *Biochimica et Biophysica Acta* **1361,** 92-102.

Gradoni, L., Pozio, E., Gramiccia, M., Maroli, M. and Bettini, S. (1983) Leishmaniasis in Tuscany (Italy): VII. Studies on the role of the black rat, *Rattus rattus*, in the epidemiology of visceral leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **77,** 427-431.

Graham, S. V. (1995) Mechanisms of stage-regulated gene expression in kinetoplastida. *Parasitology Today* **11,** 217-223.

Grimaldi, G., Jr. and Tesh, R. B. (1993) Leishmaniases of the New World: current concepts and implications for future research. *Clinical Microbiology Reviews* **6,** 230-250.

Guan, L., Yang, Y., Xu, Y., Qu, J., Zho, X., Wang, G., Lu, H., Zhong, L. and Chang, K. P. (1994) [Leishmaniasis in Karamay. XIV. Identification of promastigote isolates from naturally infected *Phlebotomus major*]. *Chung Kuo Chi Sheng Chung Hsueh Yu Chi Sheng Chung Ping Tsa Chih* **12,** 257-261.

Guevara, P., Alonso, G., Da Silveira, J. F., de Mello, M., Scorza, J. V., Anez, N. and

Ramirez, J. L. (1992) Identification of new world *Leishmania* using ribosomal gene spacer probes. *Molecular and Biochemical Parasitology* **56,** 15-26.

Guizani, I., Van Eys, G. J., Ismail, R. B. and Dellagi, K. (1994) Use of recombinant DNA probes for species identification of Old World *Leishmania* isolates. *American Journal of Tropical Medicine and Hygiene* **50,** 632-640.

Gurunathan, S., Sacks, D. L., Brown, D. R., Reiner, S. L., Charest, H., Glaichenhaus, N. and Seder, R. A. (1997) Vaccination with DNA encoding the immunodominant LACK parasite antigen confers protective immunity to mice infected with *Leishmania major*. *Journal of Experimental Medicine* **186,** 1137-1147.

Hacia, J. G., Makalowski, W., Edgemon, K., Erdos, M. R., Robbins, C. M., Fodor, S. P., Brody, L. C. and Collins, F. S. (1998) Evolutionary sequence comparisons using high-density oligonucleotide arrays. *Nature Genetics* **18,** 155-158.

Hall, L. R. and Titus, R. G. (1995) Sand fly vector saliva selectively modulates macrophage functions that inhibit killing of *Leishmania major* and nitric oxide production. *Journal of Immunology* **155,** 3501-3506.

Harris, H. and Hopkinson, D. A. (1976) *Handbook of enzyme electrophoresis in human genetics.* Amsterdam, North Holland.

Harris, E., Kropp, G., Belli, A., Rodriguez, B. and Agabian, N. (1998) Single-step multiplex PCR assay for characterization of New World *Leishmania* complexes. *Journal of Clinical Microbiology* **36,** 1989-1995.

Hassan, M. Q., Ghosh, A., Ghosh, S. S., Gupta, M., Basu, D., Mallik, K. K. and Adhya, S. (1993) Enzymatic amplification of mini-exon-derived RNA gene spacers of *Leishmania donovani*: primers and probes for DNA diagnosis. *Parasitology* **107,** 509-517.

Hausmann, K., Hulsmann, N. (1996) *Protozoology*. 2nd edition. Giorg Thieme Verlag, Inc., Sttutgard and New York

Heid, C. A., Stevens, J., Livak, K. J. and Williams, P. M. (1996) Real time quantitative PCR. *Genome Research* **6,** 986-994.

Hennig, W. (1965) Phylogenetic systematics. *Annual Review of Entomology* **10,** 97-106.

Hernandez Montes, O., Monroy Ostria, A., McCann, S. and Barker, D. C. (1998) Identification of Mexican *Leishmania* species by analysis of PCR amplified DNA. *Acta Tropica* **71,** 139-153.

Hervas, J., Mendez, A., Carrasco, L. and Gomez Villamandos, J. C. (1996) Pathological study of visceral leishmaniasis in a jackal (*Canis aureus*) *Veterinary Record* **139,** 293-295.

Hey, A. S., Theander, T. G., Hviid, L., Hazrati, S. M., Kemp, M. and Kharazmi, A. (1994) The major surface glycoprotein (gp63) from *Leishmania major* and *Leishmania donovani* cleaves CD4 molecules on human T cells. *Journal of Immunology* **152,** 4542-4548.

Higgins, D.G., Bleasby, A.J. and Fuchs, R. (1992) Clustal V: improved software for multiple alignment. *Computer Applications in the BioSciences* **8**, 189-191.

Holloway, J. W., Beghe, B., Turner, S., Hinks, L. J., Day, I. N. and Howell, W. M. (1999) Comparison of three methods for single nucleotide polymorphism typing for DNA bank studies: sequence-specific oligonucleotide probe hybridisation, TaqMan liquid phase hybridisation, and microplate array diagonal gel electrophoresis (MADGE) *Human Mutation* **14,** 340-347.

Hommel, M., Peters, W., Ranque, J., Quilici, M. and Lanotte, G. (1978) The micro-ELISA technique in the serodiagnosis of visceral leishmaniasis. *Annals of Tropical Medicine and Parasitology* **72,** 213-218.

Howard, M. K., Kelly, J. M., Lane, R. P. and Miles, M. A. (1991) A sensitive repetitive DNA probe that is specific to the *Leishmania donovani* complex and its use as an epidemiological and diagnostic reagent. *Molecular and Biochemical Parasitology* **44,** 63-72.

Hu, X., Lu, H., Luo, P., Wang, Z., Hu, X. and Yi, T. (1992) Identification of *Leishmania donovani* isolates from different kala-azar foci in China by kDNA hybridization. *Chinese Medical Sciences Journal* **7,** 63-66.

Hyams, K. C., Hanson, K., Wignall, F. S., Escamilla, J. and Oldfield, E. C.(1995) The impact of infectious diseases on the health of U.S. troops deployed to the Persian Gulf during operations Desert Shield and Desert Storm. *Clinical Infectious Diseases* **20,** 1497-1504.

Ibrahim, E. A., al Zahrani, M. A., al Tuwaigri, A. S., al Shammary, F. J. and Evans, D. A. (1992) *Leishmania* infecting man and wild animals in Saudi Arabia. 9. The black rat (*Rattus rattus*) a probable reservoir of visceral leishmaniasis in Gizan province, south-west Saudi Arabia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86,** 513-514.

Jaffe, C. L., Bennett, E., Grimaldi, G., Jr. and McMahon Pratt, D. (1984) Production and characterization of species-specific monoclonal antibodies against *Leishmania donovani* for immunodiagnosis. *Journal of Immunology* **133,** 440-447.

Jarry, D. M. (1999) Historique des leishmanioses et de leurs complexes pathogenes. (Ed. Dedet, J.-P.) Ellipses, pp. 13-20.

Jeronimo, S. M., Oliveira, R. M., Mackay, S., Costa, R. M., Sweet, J., Nascimento, E. T., Luz, K. G., Fernandes, M. Z., Jernigan, J. and Pearson, R. D. (1994) An urban outbreak of visceral leishmaniasis in Natal, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **88,** 386-368.

Joshi, P. B., Sacks, D. L., Modi, G. and McMaster, W. R. (1998) Targeted gene deletion of *Leishmania major* genes encoding developmental stage-specific leishmanolysin (GP63) *Molecular Microbiology* **27,** 519-530.

Kar, K. (1995) Serodiagnosis of leishmaniasis. *Critical Reviews in Microbiology* **21,** 123-152.

Kar, K., Mukerji, K., Naskar, K., Bhattacharya, A. and Ghosh, D. K. (1990) *Leishmania donovani*: a chemically defined medium suitable for cultivation and cloning of

promastigotes and transformation of amastigotes to promastigotes. *Journal of Protozoology* **37,** 277-279.

Karp, C. L., El Safi, S. H., Wynn, T. A., Satti, M. M., Kordofani, A. M., Hashim, F. A., Hag Ali, M., Neva, F. A., Nutman, T. B. and Sacks, D. L. (1993) In vivo cytokine profiles in patients with kala-azar. Marked elevation of both interleukin-10 and interferon-gamma. *Journal of Clinical Investigation* **91,** 1644-1648.

Katakura, K., Kawazu, S., Naya, T., Nagakura, K., Ito, M., Aikawa, M., Qu, J. Q., Guan, L. R., Zuo, X. P., Chai, J. J., Chang, K. P. and Matsumoto, Y. (1998) Diagnosis of kala-azar by nested PCR based on amplification of the *Leishmania* mini-exon gene. *Journal of Clinical Microbiology* **36,** 2173-2177.

Kaul, S. M., Sharma, R. S. and Borgohain, B. K. (1994) D.D.T. resistance in *Sergentomyia shorttii* (Diptera: Psychodidae) in Kamrup, Assam - first report in *Sergentomyia* genus. *Journal of Communicable Diseases* **26,** 100-102.

Kaye, P. M., Gorak, P., Murphy, M. and Ross, S. (1995) Strategies for immune intervention in visceral leishmaniasis. *Annals Tropical Medicine and Parasitology* **89 S1,** 75-81.

Keithly, J. S. (1989) Ornithine decarboxylase and trypanothione reductase genes in *Leishmania braziliensis guyanensis*. *Journal of Protozoology* **36,** 498-501.

Kelly, J. M., Law, J. M., Chapman, C. J., Van Eys, G. J. and Evans, D. A. (1991) Evidence of genetic recombination in *Leishmania*. *Molecular and Biochemical Parasitology* **46,** 253-263.

Kelly, J. M. (1993) Isolation of RNA and DNA from *Leishmania*. In *Methods in Molecular Biology* (Ed. Hyde, J. E.), Humana Press, pp. 123-131.

Kemp, M., Hey, A. S., Kurtzhals, J. A., Christensen, C. B., Gaafar, A., Mustafa, M. D., Kordofani, A. A., Ismail, A., Kharazmi, A. and Theander, T. G. (1994) Dichotomy of the human T cell response to *Leishmania* antigens. I. Th1-like response to *Leishmania major* promastigote antigens in individuals recovered from cutaneous leishmaniasis. *Clinical and Experimental Immunology* **96,** 410-415.

Kerr, S.F. (2000) Palearctic origin of *Leishmania*. *Memórias do Instituto Oswaldo Cruz* **95**, 75-80

Kien Truong, T., Ambroise Thomas, P., Quilici, M., Dunan, S. and Ranque, P. (1969) Diagnostic serologique des leishmanioses par immuno-fluorescence sur coupes de foies ou de rates de hamsters infestés avec *Leishmania donovani*. *Bulletin de la Societe de Pathologie Exotique et de ses Filiales* **62,** 1077-1084.

Killick Kendrick, R., Molyneux, D. H., Rioux, J. A., Lanotte, G. and Leaney, A. J. (1980) Possible origins of *Leishmania chagasi*. *Annals of Tropical Medicine and Parasitology* **74,** 563-5.

Killick Kendrick, R. and Rioux, J. A. (1991) Intravectorial cycle of *Leishmania* in sandflies. *Annales de Parasitologie Humaine et Comparee* **66,** 71-74.

Killick Kendrick, R., Lainson, R., Rioux, J. A. and Saf'janova, V.M. (1986) The taxonomy of *Leishmania*-like parasites of reptiles. In *Leishmania. Taxonomie et phylogenese.*

Applications eco-epidemiologiques. (Ed. Rioux, J-A) Montpellier, pp. 143-148.

Kluge, A.G. and Farris, J.S. (1969) Quantitative phyletics and the evolution of anurans. *Systematic Zoology.* **18**, 1-32.

Kreutzer, R. D., Yemma, J. J., Grogl, M., Tesh, R. B. and Martin, T. I. (1994) Evidence of sexual reproduction in the protozoan parasite *Leishmania* (Kinetoplastida: Trypanosomatidae). *American Journal of Tropical Medicine and Hygiene* **51,** 301-307.

Kurtzhals, J. A., Hey, A. S., Jardim, A., Kemp, M., Schaefer, K. U., Odera, E. O., Christensen, C. B., Githure, J. I., Olafson, R. W., Theander, T. G. and Kharazmi, A. (1994) Dichotomy of the human T cell response to *Leishmania* antigens. II. Absent or Th2-like response to gp63 and Th1-like response to lipophosphoglycan-associated protein in cells from cured visceral leishmaniasis patients. *Clinical and Experimental Immunology* **96,** 416-421.

Lainson, R., Dye, C., Shaw, J. J., Macdonald, D. W., Courtenay, O., Souza, A. A. and Silveira, F. T. (1990) Amazonian visceral leishmaniasis--distribution of the vector *Lutzomyia longipalpis* (Lutz & Neiva) in relation to the fox *Cerdocyon thous* (Linn.) and the efficiency of this reservoir host as a source of infection. *Memórias do Instituto Oswaldo Cruz* **85,** 135-137.

Lainson, R. and Shaw, J. J. (1972) Leishmaniasis of the New World: taxonomic problems. *British Medical Bulletin* **28,** 44-48.

Lainson, R. and Shaw, J. J. (1979) The role of animals in the epidemiology of South American leishmaniasis. In *Biology of the Kinetoplastida* (Eds. Lumsden, W. H. R. and Evans, D. A.) Academic Press, pp. 1-116.

Lainson, R. and Shaw, J. J. (1987) Evolution, classification and geographical distribution. In *The leishmaniasis in biology and medicine* (Eds. Peters, W. and Killick-Kendrick, R.) Academic Press, pp 1-120.

Lainson, R. and Shaw, J. J. (1998) New World Leishmaniasis - The neotropical *Leishmania* species. In *Parasitology* (Eds. Cox, F. E. G., Kreier, J. P. and Wakelin, D.) Arnold, pp. 241-266.

Lainson, R., Shaw, J. J. and Lins, Z. C. (1969) Leishmaniasis in Brazil. IV. The fox, *Cerdocyon thous* (L) as a reservoir of *Leishmania donovani* in Para state, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **63,** 741-745.

Lainson, R., Shaw, J. J., Silveira, F. T. and Braga, R. R. (1987) American visceral leishmaniasis: on the origin of *Leishmania* (*Leishmania*) *chagasi. Transactions of the Royal Society of Tropical Medicine and Hygiene* **81,** 517.

Lamboy, W. F. (1994) Computing genetic similarity coefficients from RAPD data: the effects of PCR artefacts. *PCR Methods and Applications* **4,** 31-37.

Lang, T., Warburg, A., Sacks, D. L., Croft, S. L., Lane, R. P. and Blackwell, J. M. (1991) Transmission and scanning EM-immunogold labelling of *Leishmania major* lipophosphoglycan in the sandfly *Phlebotomus papatasi. European Journal of Cell Biology* **55,** 362-372.

Lanotte, G. and Rioux, J. A. (1990) Fusion cellulaire chez les *Leishmania* (Kinetoplastida, Trypanosomatidae). *Comptes Rendus de l'Academie des Sciences Serie III* **310,** 285-288.

Laskay, T., Gemetchu, T., Teferedegn, H. and Frommel, D. (1991) The use of DNA hybridization for the detection of *Leishmania aethiopica* in naturally infected sandfly vectors. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **85,** 599-602.

Launois, P., Maillard, I., Pingel, S., Swihart, K. G., Xenarios, I., Acha Orbea, H., Diggelmann, H., Locksley, R. M., MacDonald, H. R. and Louis, J. A. (1997) IL-4 rapidly produced by V beta 4 V alpha 8 CD4+ T cells instructs Th2 development and susceptibility to *Leishmania major* in BALB/c mice. *Immunity* **6,** 541-549.

Le Blancq, S. M. (1986) An epidemiological and taxonomical study of Old World *Leishmania* using isoenzymes. *Molecular Parasitology Department*, University of London, 286 pp.

Le Blancq, S. M., Cibulskis, R. E. and Peters, W. (1986) *Leishmania* in the Old World: 5. Numerical analysis of isoenzyme data. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **80,** 517-524.

Le Blancq, S. M. and Peters, W. (1986) *Leishmania* in the Old World: 4. The distribution of *L. donovani sensu lato* zymodemes. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **80,** 367-377.

Le Fichoux, Y., Mary, C., Marty, P. and Kubar, J. (1999) Diagnostic des leishmanioses. In *Les leishmanioses* (*Ed.* Dedet, J. P.), Ellipses, pp. 191-203.

LeFurgey, A., Ingram, P. and Blum, J. J. (1990) Elemental composition of polyphosphate-containing vacuoles and cytoplasm of *Leishmania major*. *Molecular and Biochemical Parasitology* **40,** 77-86.

Lerner, E. A. and Shoemaker, C. B. (1992) Maxadilan. Cloning and functional expression of the gene encoding this potent vasodilator peptide. *Journal of Biological Chemistry* **267,** 1062-1066.

Lewis, D.J. (1982) A taxonomic review of the genus *Phlebotomus* (Diptera: Psychodidae). *Bulletin of the British Museum of Natural History* **45**, 121-209

Lighthall, G. K. and Giannini, S. H. (1992) The chromosomes of *Leishmania*. *Parasitology Today* **8,** 192-199.

Lipsitch, M. and Moxon, E. R. (1997) Virulence and transmissibility of pathogens: what is the relationship? *Trends in Microbiology* **5,** 31-37.

Louis, J., Himmelrich, H., Parra Lopez, C., Tacchini Cottier, F. and Launois, P. (1998) Regulation of protective immunity against *Leishmania major* in mice. *Current Opinion in Immunology* **10,** 459-464.

Lu, H. G. and Hu, X. S. (1990) Identification of *Leishmania* by kinetoplast DNA minicircle and cloning of minicircle. *Chinese Medical Journal* **103,** 418-423.

Lu, H. G., Zhong, L., Chang, K. P. and Docampo, R. (1997) Intracellular Ca2+ pool content

and signaling and expression of a calcium pump are linked to virulence in *Leishmania mexicana amazonesis* amastigotes. *Journal of Biological Chemistry* **272,** 9464-9473.

Lupascu, G., Iancu, L., Bossie, A. and Ciurea, C. (1970) Valeur de la reaction d'immunofluorescence pour le depistage des infections a *P. malariae* et la dynamique des titres des anticorps au cours de l'infection et apres le traitement radical. *Archives Roumaines de Pathologie Experimentale et de Microbiologie* **29,** 281-287.

Macedo, A. M., Melo, M. N., Gomes, R. F. and Pena, S. D. (1992) DNA fingerprints: a tool for identification and determination of the relationships between species and strains of *Leishmania*. *Molecular and Biochemical Parasitology* **53,** 63-70.

Maingon, R. D. C., Elson, L., Hommel, M., Crampton, J. M. and Chance, M. L. (1990) Multiple distinct gp63-like genes in amastigotes of *Leishmania donovani*, In *Parasites: molecular biology, drug and vaccine design*. Proceedings of a UCLA Symposium held at Keystone, Colorado, April 3-10, 1989. Willey-Liss Inc., New York. (Ed. Agabian, N. and Cerami, A.) pp. 191-204.

Maingon, R., Feliciangeli, D., Ward, R., Chance, M., Adamson, R., Rodriguez, N., Convit, J., Petralanda, I., Hernandez, A. and Segovia, M. (1993) Molecular approaches applied to the epidemiology of leishmaniasis in Venezuela. *Archives de l'Institut Pasteur de Tunis* **70,** 309-324.

Marras, S. A., Kramer, F. R. and Tyagi, S. (1999) Multiplex detection of single-nucleotide variations using molecular beacons. *Genetic Analysis* **14,** 151-156.

Marty, P., Le Fichoux, Y., Giordana, D. and Brugnetti, A. (1992) Leishmanin reaction in the human population of a highly endemic focus of canine leishmaniasis in Alpes-Maritimes, France. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86,** 249-250.

Marty, P., Lelievre, A., Quaranta, J. F., Rahal, A., Gari Toussaint, M. and Le Fichoux, Y. (1994) Use of the leishmanin skin test and western blot analysis for epidemiological studies in visceral leishmaniasis areas: experience in a highly endemic focus in Alpes-Maritimes (France) *Transactions of the Royal Society of Tropical Medicine and Hygiene* **88,** 658-659.

Massamba, N. N. and Mutinga, M. J. (1992) Recombinant kinetoplast DNA (kDNA) probe for identifying *Leishmania tropica*. *Acta Tropica* **52,** 1-15.

Mauricio, I., Howard, M., Stothard, J. and Miles, M. (1999) Genetic diversity in the *Leishmania donovani* complex. *Parasitology* **119,** 237-246.

Mauricio, I. L., Stothard, J. R. and Miles, M. A. (2000) The strange case of *Leishmania chagasi*. *Parasitology Today* **16,** 188-199.

Maxam, A. M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proceedings of the National Academy of Sciences USA* **74,** 560-564.

Maxam, A. M. and Gilbert, W. (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods in Enzymology* **65,** 499-560.

May, B. (1992) Starch gel electrophoresis of allozymes in *Molecular genetic analysis of*

populations. A practical approach. (Ed. Hoelzel), Oxford University Press, pp. 1-28.

Maynard Smith, J., O'Rourke, M. and Spratt, B. G. (1993) How clonal are bacteria? *Proceedings of the National Academy of Sciences USA* **90,** 4384-4388.

Mayrink, W., Chiari, C. d. A., Magalhaes, P. A. and Costa, C. A. d. (1972) Teste do latex no diagnóstico do calazar americano. *Revista do Instituto de Medicina Tropical de São Paulo* **14,** 273-276.

Mayrink, W., Genaro, O., Silva, J. C., Da Costa, R. T., Tafuri, W. L., Toledo, V. P., Da Silva, A. R., Reis, A. B., Williams, P. and Da Costa, P. W. (1996) Phase I and II open clinical trials of a vaccine against *Leishmania chagasi* infections in dogs. *Memórias do Instituto Oswaldo Cruz* **91,** 695-697.

Mbow, M. L., Bleyenberg, J. A., Hall, L. R. and Titus, R. G. (1998) *Phlebotomus papatasi* sand fly salivary gland lysate down-regulates a Th1, but up-regulates a Th2, response in mice infected with *Leishmania major*. *Journal of Immunology* **161,** 5571-5577.

McMahon Pratt, D. and David, J. R. (1981) Monoclonal antibodies that distinguish between New World species of *Leishmania*. *Nature* **291,** 581-583.

Meade, J. C., Hudson, K. M., Stringer, S. L. and Stringer, J. R. (1989) A tandem pair of *Leishmania donovani* cation transporting ATPase genes encode isoforms that are differentially expressed. *Molecular and Biochemical Parasitology* **33,** 81-91.

Mebrahtu, Y. B., Lawyer, P. G., Hendricks, L. D., Muigai, R., Oster, C. N., Perkins, P. V., Koech, D. K., Pamba, H. and Roberts, C. R. (1991) Concurrent infection with *Leishmania donovani* and *Leishmania major* in a Kenyan patient: clinical description and parasite characterization. *American Journal of Tropical Medicine and Hygiene* **45,** 290-296.

Mebrahtu, Y. B., Lawyer, P. G., Pamba, H., Koech, D., Perkins, P. V., Roberts, C. R., Were, J. B. and Hendricks, L. D. (1992) Biochemical characterization and zymodeme classification of *Leishmania* isolates from patients, vectors, and reservoir hosts in Kenya. *American Journal of Tropical Medicine and Hygiene* **47,** 852-892.

Mebrahtu, Y. B., Van Eys, G., Guizani, I., Lawyer, P. G., Pamba, H., Koech, D., Roberts, C., Perkins, P. V., Were, J. B. and Hendricks, L. D. (1993) Human cutaneous leishmaniasis caused by *Leishmania donovani s.l.* in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **87,** 598-601.

Medina Acosta, E., Karess, R. E. and Russell, D. G. (1993) Structurally distinct genes for the surface protease of *Leishmania mexicana* are developmentally regulated. *Molecular and Biochemical Parasitology* **57,** 31-45.

Medina Acosta, E., Karess, R. E., Schwartz, H. and Russell, D. G. (1989) The promastigote surface protease (gp63) of *Leishmania* is expressed but differentially processed and localised in the amastigote stage. *Molecular and Biochemical Parasitology* **37,** 263-273.

Mello, D. A., Rego Junior, F. d. A., Oshozo, E. and Nunes, V. L. (1989) *Cerdocyon thous* (L.) (Carnivora, Canidae) naturally infected with *Leishmania donovani chagasi* (Cunha

& Chagas, 1973) in Corumba (Mato Grosso do Sul State, Brazil). *Memórias do Instituto Oswaldo Cruz* **83,** 259.

Mendez, S., Nell, M. and Alunda, J. M. (1996) *Leishmania infantum*: infection of macrophages in vitro with promastigotes. *International Journal for Parasitology* **26,** 619-622.

Mendoza Leon, A., Havercroft, J. C. and Barker, D. C. (1995) The RFLP analysis of the beta-tubulin gene region in New World *Leishmania*. *Parasitology* **111,** 1-9.

Mengesha, B. and Abuhoy, M. (1978) Kala-azar among labour migrants in Metema-Humera region of Ethiopia. *Tropical and Geographical Medicine* **30,** 199-206.

Millesimo, M., Zucca, M., Caramello, P. and Savoia, D. (1996) Evaluation of the immune response in visceral leishmaniasis. *Diagnostic Microbiology and Infectious Disease* **26,** 7-11.

Miles, M. A., Lanham, S. M., de Souza, A. A. and Povoa, M. (1980a) Further enzymic characters of Trypanosoma cruzi and their evaluation for strain identification. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **74,** 221-37.

Miles, M. A., Povoa, M. M., de Souza, A. A., Lainson, R. and Shaw, J. J. (1980b) Some methods for the enzymic characterization of Latin-American *Leishmania* with particular reference to *Leishmania mexicana amazonensis* and subspecies of *Leishmania hertigi*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **74,** 243-52.

Modabber, F. (1990) Development of vaccines against leishmaniasis. *Scandinavian Journal of Infectious Diseases* **76S,** 72-8.

Mohammed El, A. and Wright, E. P. (1987) Monoclonal antibody identification of *Leishmania tropica* from Sudanese mucosal leishmaniasis. *Tropical Medicine and Parasitology* **38,** 34-36.

Molina, R., Amela, C., Nieto, J., San Andres, M., Gonzalez, F., Castillo, J. A., Lucientes, J. and Alvar, J. (1994) Infectivity of dogs naturally infected with *Leishmania infantum* to colonised *Phlebotomus perniciosus*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **88,** 491-493.

Molyneux, D. H. (1974) Virus-like particles in *Leishmania* parasites. *Nature* **249,** 588-589.

Molyneux, D. H. (1998) Vector-borne parasitic diseases--an overview of recent changes. *International Journal for Parasitology* **28,** 927-934.

Momen, H. and Cupollillo, E. (2000) Speculations on the origin and evolution of the genus *Leishmania*. *Memorias do Instituto Oswaldo Cruz*. **95**: 583-588.

Momen, H., Grimaldi Junior, G. and Deane, L. M. (1987) *Leishmania infantum*, the aetiological agent of American visceral leishmaniasis (AVL)? *Memórias do Instituto Oswaldo Cruz* **82,** 447-448.

Momen, H., Pacheco, R. S., Cupolillo, E. and Grimaldi Junior, G. (1993) Molecular evidence for the importation of Old World *Leishmania* into the Americas. *Biological Research* **26,** 249-255.

Moody, S. F. (1993) Molecular variation in *Leishmania*. *Acta Tropica* **53,** 185-204.

Morales, G., Carrillo, G., Requena, J. M., Guzman, F., Gomez, L. C., Patarroyo, M. E. and Alonso, C. (1997) Mapping of the antigenic determinants of the *Leishmania infantum* gp63 protein recognized by antibodies elicited during canine visceral leishmaniasis. *Parasitology* **114,** 507-516.

Moreno, G., Rioux, J.-A., Lanotte, G., Pratlong, F. and Serres, E. (1984) Le complexe *Leishmania donovani s.l.* Analyse enzymatique et traitement numerique. Individualization du complexe *Leishmania infantum*. Corollaires biogeographiques et phyletiques. Apropos de 146 souches originaires de l'Ancien et du Nouveau Monde. In *Leishmania. Taxonomie et phylogenese. Applications eco-epidemiologiques.* (Ed. Rioux, J-A) Montpellier, pp. 105-117.

Morsy, T. A., Khalil, N. M., Salama, M. M., Hamdi, K. N., al Shamrany, Y. A. and Abdalla, K. F. (1995) Mucosal leishmaniasis caused by *Leishmania tropica* in Saudi Arabia. *Journal of the Egyptian Society of Parasitology* **25,** 73-79.

Mosser, D. M. and Karp, C. L. (1999) Receptor mediated subversion of macrophage cytokine production by intracellular pathogens. *Current Opinion in Immunology* **11,** 406-411.

Motazedian, H., Noyes, H. and Maingon, R. (1996) *Leishmania* and *Sauroleishmania*: the use of random amplified polymorphic DNA for the identification of parasites from vertebrates and invertebrates. *Experimental Parasitology* **83,** 150-154.

Mukhopadhyay, A. K., Hati, A. K., Chakraborty, S. and Saxena, N. B. (1996) Effect of DDT on *Phlebotomus* sandflies in kala-azar endemic foci in West Bengal. *Journal of Communicable Diseases* **28,** 171-175.

Murphy, K. M. (1998) T lymphocyte differentiation in the periphery. *Current Opinion in Immunology* **10,** 226-232.

Murray, P. J. and Spithill, T. W. (1991) Variants of a *Leishmania* surface antigen derived from a multigenic family. *Journal of Biological Chemistry* **266,** 24477-24484.

Murray, P. J., Spithill, T. W. and Handman, E. (1989) The PSA-2 glycoprotein complex of *Leishmania major* is a glycosylphosphatidylinositol-linked promastigote surface antigen. *Journal of Immunology* **143,** 4221-4226.

Nandy, A., Neogy, A. B. and Chowdhury, A. B. (1987) Leishmanin test survey in an endemic village of Indian kala-azar near Calcutta. *Annals of Tropical Medicine and Parasitology* **81,** 693-699.

Nelson, R. G., Parsons, M., Barr, P. J., Stuart, K., Selkirk, M. and Agabian, N. (1983) Sequences homologous to the variant antigen mRNA spliced leader are located in tandem repeats and variable orphons in *Trypanosoma brucei*. *Cell* **34,** 901-909.

Nilsen, T. W. (1994) Unusual strategies of gene expression and control in parasites. *Science* **264,** 1868-1869.

Noyes, H. (1998a) Can *Trypanosoma* trees be trusted? *Parasitology Today* **14,** 49-50.

Noyes, H. (1998b) Implications of a Neotropical origin of the genus *Leishmania*. *Memórias do Instituto Oswaldo Cruz* **93,** 657-661.

Noyes, H. A., Arana, B. A., Chance, M. L. and Maingon, R. (1997) The *Leishmania hertigi* (Kinetoplastida; Trypanosomatidae) complex and the lizard *Leishmania*: their classification and evidence for a neotropical origin of the *Leishmania-Endotrypanum* clade. *Journal of Eukaryotic Microbiology* **44,** 511-517.

Noyes, H. A., Belli, A. A. and Maingon, R. (1996) Appraisal of various random amplified polymorphic DNA-polymerase chain reaction primers for *Leishmania* identification. *American Journal of Tropical Medicine and Hygiene* **55,** 98-105.

Nunes, A. C. and Ramalho Pinto, F. J. (1996) Complement resistance of *Leishmania amazonensis* promastigotes is independent of parasite proteases and lysis of sensitive forms is not due to natural antibodies in normal human serum. *Brazilian Journal of Medicine and Biology Research* **29,** 1633-1640.

Oliveira Neto, M. P., Grimaldi, G., Jr., Momen, H., Pacheco, R. S., Marzochi, M. C. and McMahon Pratt, D. (1986a) Active cutaneous leishmaniasis in Brazil, induced by *Leishmania donovani chagasi*. *Memórias do Instituto Oswaldo Cruz* **81,** 303-309.

Oliveira Neto, M. P., Marzochi, M. C., Grimaldi Junior, G., Pacheco, R. S., Toledo, L. M. and Momen, H. (1986b) Concurrent human infection with *Leishmania donovani* and *Leishmania braziliensis braziliensis*. *Annals of Tropical Medicine and Parasitology* **80,** 587-592.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences USA* **86,** 2766-2770.

Oskam, L., Pratlong, F., Zijlstra, E. E., Kroon, C. C., Dedet, I. P., Kager, P. A., Schonian, G., Ghalib, H. W., El Hassan, A. M. and Meredith, S. E. (1998) Biochemical and molecular characterization of *Leishmania* parasites isolated from an endemic focus in eastern Sudan. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92,** 120-122.

Ozon, C., Marty, P., Pratlong, F., Breton, C., Blein, M., Lelievre, A. and Haas, P. (1998) Disseminated feline leishmaniosis due to *Leishmania infantum* in Southern France. *Veterinary Parasitology* **75,** 273-277.

Page, R. D. M. and Holmes, E. C. (1998) *Molecular evolution. A phylogenetic approach.* Blackwell Science Ltd., Oxford. 346 pp.

Palatnik, C. B., Previato, J. O., Mendonca Previato, L. and Borojevic, R. (1990) A new approach to the phylogeny of *Leishmania*: species specificity of glycoconjugate ligands for promastigote internalisation into murine macrophages. *Parasitology Research* **76,** 289-293.

Palma, G. and Gutierrez, Y. (1991) Laboratory diagnosis of *Leishmania*. *Clinics in Laboratory Medicine* **11,** 909-922.

Pan, A. A., Duboise, S. M., Eperon, S., Rivas, L., Hodgkinson, V., Traub Cseko, Y. and McMahon Pratt, D. (1993) Developmental life cycle of *Leishmania* - cultivation and

characterization of cultured extracellular amastigotes. *Journal of Eukaryotic Microbiology* **40,** 213-223.

Panton, L. J., Tesh, R. B., Nadeau, K. C. and Beverley, S. M. (1991) A test for genetic exchange in mixed infections of *Leishmania major* in the sand fly *Phlebotomus papatasi*. *Journal of Protozoology* **38,** 224-228.

Pappas, M. G., Hajkowski, R., Diggs, C. L. and Hockmeyer, W. T. (1983) Development of an antigen conservative enzyme immunoassay (Dot-ELISA) for the rapid diagnosis of visceral leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **77,** 425-426.

Pascale, J. M., Sousa, O. E. and Lopez, U. (1992) Clonaje de minicirculos de kADN en diferentes especies de *Leishmania* y su uso como sondas para diagnostico. *Revista Medica de Panama* **17,** 163-172.

Piarroux, R., Fontes, M., Perasso, R., Gambarelli, F., Joblet, C., Dumon, H. and Quilici, M. (1995) Phylogenetic relationships between Old World *Leishmania* strains revealed by analysis of a repetitive DNA sequence. *Molecular and Biochemical Parasitology* **73,** 249-252.

Pimenta, P. F., Pinto Da Silva, P., Rangarajan, D., Smith, D. F. and Sacks, D. L. (1994) *Leishmania major*: association of the differentially expressed gene B protein and the surface lipophosphoglycan as revealed by membrane capping. *Experimental Parasitology* **79,** 468-479.

Pinelli, E., Boog, C. J., Rutten, V. P., van Dijk, B., Bernadina, W. E. and Ruitenberg, E. J. (1994) A canine CD8+ cytotoxic T-cell line specific for *Leishmania infantum*-infected macrophages. *Tissue Antigens* **43,** 189-192.

Podani (1993) SYN-TAX-pc Computer Programs for Multivariate Data Analysis in Ecology and Systematics, Sciencia Publishing.

Ponce, C., Ponce, E., Morrison, A., Cruz, A., Kreutzer, R., McMahon Pratt, D. and Neva, F. (1991) *Leishmania donovani chagasi*: new clinical variant of cutaneous leishmaniasis in Honduras. *Lancet* **337,** 67-70.

Pratlong, F. and Lanotte, G. (1999) Identification, taxonomie et phylogenese. In *Les leishmanioses* (Ed. Dedet, J.-P.) Ellipses, pp. 21-39.

Pratlong, F., Portus, M., Rispail, P., Moreno, G., Bastien, P. and Rioux, J. A. (1989) Presence simultanee chez le chien de deux zymodemes du complexe *Leishmania infantum*. *Annales de Parasitologie Humaine et Comparee* **64,** 312-314.

Pratlong, F., Roux, M., Rispail, P., Lannote, G. and Rioux, J. A. (1986) Utilization du pouvoir pathogene experimental comme critere taxonomique du genre *Leishmania* Ross, 1903. Analyse critique a propos d'une experimentation portant sur 55 souches de l'Ancien Monde. In *Leishmania. Taxonomie et phylogenese. Applications eco-epidemiologiques.* (Ed. Rioux, J-A) Montpellier, pp. 199-210.

Qiao, Z., Miles, M. A. and Wilson, S. M. (1995) Detection of parasites of the *Leishmania donovani*-complex by a polymerase chain reaction-solution hybridization enzyme-

linked immunoassay (PCR-SHELA) *Parasitology* **110,** 269-275.

Ramamoorthy, R., Donelson, J. E., Paetz, K. E., Maybodi, M., Roberts, S. C. and Wilson, M. E. (1992) Three distinct RNAs for the surface protease gp63 are differentially expressed during development of *Leishmania donovani chagasi* promastigotes to an infectious form. *Journal of Biological Chemistry* **267,** 1888-1895.

Ramesh, V. (1994) On the differences between post kala-azar dermal leishmaniasis and leprosy. *Tropical Doctor* **24,** 120-121.

Ramirez, J. L. and Guevara, P. (1987) The ribosomal gene spacer as a tool for the taxonomy of *Leishmania*. *Molecular and Biochemical Parasitology* **22,** 177-183.

Ramos, A., Maslov, D. A., Fernandes, O., Campbell, D. A. and Simpson, L. (1996) Detection and identification of human pathogenic *Leishmania* and *Trypanosoma* species by hybridization of PCR-amplified mini-exon repeats. *Experimental Parasitology* **82,** 242-250.

Ramsay, G. (1998) DNA chips: state-of-the art. *Nature Biotechnology* **16,** 40-4.

Ravel, C., Wincker, P., Bastien, P., Blaineau, C. and Pages, M. (1995) A polymorphic minisatellite sequence in the subtelomeric regions of chromosomes I and V in *Leishmania infantum*. *Molecular and Biochemical Parasitology* **74,** 31-41.

Ravine, D. (1999) Automated mutation analysis. *Journal of Inherited Metabolic Diseases* **22,** 503-518.

Refaeli, Y. and Abbas, A. K. (1998) Role of cytokines in autoimmunity. *European Cytokine Network* **9,** 81-82.

Reiner, S. L. and Locksley, R. M. (1995) The regulation of immunity to *Leishmania major*. *Annual Reviews of Immunology* **13,** 151-177.

Riesner, D., Steger, G., Zimmat, R., Owens, R. A., Wagenhofer, M., Hillen, W., Vollbach, S. and Henco, K. (1989) Temperature-gradient gel electrophoresis of nucleic acids: analysis of conformational transitions, sequence variations, and protein-nucleic acid interactions. *Electrophoresis* **10,** 377-389.

Rioux, J. A., Albaret, J. L., Houin, R., Dedet, J. P. and Lanotte, G. (1968) Ecologie des leishmanioses dans le sud de la France. 2. Les reservoirs selvatiques. Infestation spontanee du Renard (*Vulpes vulpes* L.). *Annales de Parasitologie Humaine et Comparee* **43,** 421-428.

Rioux, J. A., Lanotte, G., Croset, H. and Dedet, J. P. (1972) Ecologie des leishmaniose dans le sud de la France. 5. Pouvoir infestant compare des diverses formes de leishmaniose canine vis-a-vis de *Phlebotomus ariasi* Tonnoir, 1921. *Annales de Parasitologie Humaine et Comparee* **47,** 413-419.

Rioux, J. A., Lanotte, G., Serres, E., Pratlong, F., Bastien, P. and Perieres, J. (1990) Taxonomy of *Leishmania.* Use of isoenzymes. Suggestions for a new classification. *Annales de Parasitologie Humaine et Comparee* **65,** 111-125.

Roberts, S. C., Swihart, K. G., Agey, M. W., Ramamoorthy, R., Wilson, M. E. and Donelson, J. E. (1993) Sequence diversity and organization of the *msp* gene family

encoding gp63 of *Leishmania chagasi*. *Molecular and Biochemical Parasitology* **62,** 157-171.

Rodgers, M. R., Popper, S. J. and Wirth, D. F. (1990) Amplification of kinetoplast DNA as a tool in the detection and diagnosis of *Leishmania*. *Experimental Parasitology* **71,** 267-75.

Rodriguez Cuartero, A. and Nunez Carril, J. (1974) Las proteinas plasmaticas en el Kala-Azar. Estudio evolutivo de dos casos mediante tecnicas de inmunodifusion. *Revista Clinica Espanola* **134,** 467-471.

Rodriguez, N., Guzman, B., Rodas, A., Takiff, H., Bloom, B. R. and Convit, J. (1994) Diagnosis of cutaneous leishmaniasis and species discrimination of parasites by PCR and hybridization. *Journal of Clinical Microbiology* **32,** 2246-2252.

Rossi, V., Wincker, P., Ravel, C., Blaineau, C., Pages, M. and Bastien, P. (1994) Structural organisation of microsatellite families in the *Leishmania* genome and polymorphisms at two (CA)n loci. *Molecular and Biochemical Parasitology* **65,** 271-282.

Rudenko, G. (1999) Mechanisms mediating antigenic variation in *Trypanosoma brucei*. *Memórias do Instituto Oswaldo Cruz* **94,** 235-227.

Russell, D. G. (1987) The macrophage-attachment glycoprotein gp63 is the predominant C3-acceptor site on *Leishmania mexicana* promastigotes. *European Journal of Biochemistry* **164,** 213-221.

Russell, D. G. (1995) Of microbes and macrophages: entry, survival and persistence. *Current Opinion in Immunology* **7,** 479-484.

Russell, D. G., Xu, S. and Chakrabarty, P. (1992) Intracellular trafficking and the parasitophorous vacuole of *Leishmania mexicana*-infected macrophages. *Journal of Cell Science* **103,** 1193-1210.

Russell, R., Iribar, M. P., Lambson, B., Brewster, S., Blackwell, J. M., Dye, C. and Ajioka, J. W. (1999) Intra and inter-specific microsatellite variation in the *Leishmania* subgenus *Viannia*. *Molecular and Biochemical Parasitology* **103,** 71-77.

Russo, D. M., Burns, J. M., Jr., Carvalho, E. M., Armitage, R. J., Grabstein, K. H., Button, L. L., McMaster, W. R. and Reed, S. G. (1991) Human T cell responses to gp63, a surface antigen of *Leishmania*. *Journal of Immunology* **147,** 3575-3580.

Russo, D. M., Jardim, A., Carvalho, E. M., Sleath, P. R., Armitage, R. J., Olafson, R. W. and Reed, S. G. (1993) Mapping human T cell epitopes in *Leishmania* gp63. Identification of cross-reactive and species-specific epitopes. *Journal of Immunology* **150,** 932-99.

Sadlova, J. and Volf, P. (1999) Occurrence of *Leishmania major* in sandfly urine. *Parasitology* **118,** 455-460.

Saitou, N. and Nei, M. (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. **4**: 406-425.

Sampaio-Silva, M. L., Abranches, P., Campino, L., Avelino, I. C., Santos-Gomes, G. M. and Seixas-Lopes, A. (1993) Kala-azar em Portugal. VIII. Estudo epidemiológico da

leishmaniose canina em três concelhos do Alto Douro. *Revista Portuguesa de Doenças Infecciosas* **14,** 241-243.

Samuelson, J., Lerner, E., Tesh, R. and Titus, R. (1991) A mouse model of *Leishmania braziliensis braziliensis* infection produced by coinjection with sandfly saliva. *Journal of Experimental Medicine* **173,** 49-54.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* **74,** 5463-5467.

Santoro, F., L., L. J., Rizvi, F. S., Afchain, D. and Sadigursky, M. (1986) Specificité au niveau des proteines de surface des promastigotes de *Leishmania donovani* (Laveran et Mesnil, 1903), *Leishmania infantum* Nicolle, 1908 et *Leishmania chagasi* Cunha et Chagas, 1937. In *Leishmania. Taxonomie et phylogenese. Applications eco-epidemiologiques.* (Ed. Rioux, J-A) Montpellier, pp. 71-75.

Saxena, N. B., Aggarwal, V., Dhillon, G. P., Sharma, R. S. and Rao, J. S. (1996) Visceral leishmaniasis control in India through primary health care system - a successful experiment of district level planning. *Journal of Communicable Diseases* **28,** 122-128.

Scharton Kersten, T. M. and Sher, A. (1997) Role of natural killer cells in innate resistance to protozoan infections. *Current Opinion in Immunology* **9,** 44-51.

Schnur, L. F., Zuckerman, A. and Greenblatt, C. L. (1972) Leishmanial serotypes as distinguished by the gel diffusion of factors excreted in vitro and in vivo. *Israel Journal of Medical Sciences* **8,** 932-942.

Schonian, G., Schweynoch, C., Zlateva, K., Oskam, L., Kroon, N., Graser, Y. and Presber, W. (1996) Identification and determination of the relationships of species and strains within the genus *Leishmania* using single primers in the polymerase chain reaction. *Molecular and Biochemical Parasitology* **77,** 19-29.

Schottelius, J. (1982) Lectin binding strain-specific carbohydrates on the cell surfaces of *Leishmania* strains from the Old World. *Zeitschrift Fur Parasitenkunde* **66**, 237-247

Scott, D. A., Docampo, R., Dvorak, J. A., Shi, S. and Leapman, R. D. (1997) *In situ* compositional analysis of acidocalcisomes in *Trypanosoma cruzi*. *Journal of Biological Chemistry* **272,** 28020-28029.

Seaman, J., Mercer, A. J. and Sondorp, E. (1996) The epidemic of visceral leishmaniasis in western Upper Nile, southern Sudan: course and impact from 1984 to 1994. *International Journal of Epidemiology* **25,** 862-871.

Semião Santos, S. J., El Harith, A., Ferreira, E., Pires, C. A., Sousa, C. and Gusmao, R. (1995) Evora district as a new focus for canine leishmaniasis in Portugal. *Parasitology Research* **81,** 235-259.

Shames, S. L., Fairlamb, A. H., Cerami, A. and Walsh, C. T. (1986) Purification and characterization of trypanothione reductase from *Crithidia fasciculata*, a newly discovered member of the family of disulphide-containing flavoprotein reductases. *Biochemistry* **25,** 3519-3526.

Shapiro, T. A. and Englund, P. T. (1995) The structure and replication of kinetoplast DNA.

Annual Review of Microbiology **49,** 117-143.

Shaw, J. J. (1994) Taxonomy of the genus *Leishmania*: present and future trends and their implications. *Memórias do Instituto Oswaldo Cruz* **89,** 471-478.

Shaw, J. J. and Bird, R. G. (1969) The endoerythrocytic habitat of a member of the Trypanosomatidae, *Endotrypanum schaudinni*, Mesnil and Brimont, 1908. Zeitschrift Fur Tropenmedizin Und Parasitologie **20,** 144-150.

Shiddo, S. A., Aden Mohamed, A., Akuffo, H. O., Mohamud, K. A., Herzi, A. A., Herzi Mohamed, H., Huldt, G., Nilsson, L. A., Ouchterlony, O. and Thorstensson, R. (1995) Visceral leishmaniasis in Somalia: prevalence of markers of infection and disease manifestations in a village in an endemic area. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **89,** 361-365.

Shreffler, W. G., Burns, J. M., Jr., Badaro, R., Ghalib, H. W., Button, L. L., McMaster, W. R. and Reed, S. G. (1993) Antibody responses of visceral leishmaniasis patients to gp63, a major surface glycoprotein of *Leishmania* species. *Journal of Infectious Diseases* **167,** 426-430.

Siddig, M., Ghalib, H., Shillington, D. C., Petersen, E. A. and Khidir, S. (1990) Visceral leishmaniasis in Sudan. Clinical features. *Tropical and Geographical Medicine* **42,** 107-112.

Silveira, F. T., Lainson, R., Shaw, J. J. and Povoa, M. M. (1982) Leishmaniasis in Brazil: XVIII. Further evidence incriminating the fox *Cerdocyon thous* (L) as a reservoir of Amazonian visceral leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **76,** 830-832.

Silveira, F. T., Lainson, R., Shaw, J. J. and Ribeiro, R. da S. (1984) Leishmaniose cutânea na Amazonia. Registro do primeiro caso humano de infecção mista, determinado por duas espécies distintas de leishmanias: *Leishmania brasiliensis* e *Leishmania mexicana amazonensis*. *Revista do Instituto de Medicina Tropical de São Paulo* **26,** 272-275.

Simpson, L. (1987) The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication, and evolution. *Annual Review of Microbiology* **41,** 363-382.

Singh, B. (1997) Molecular methods for diagnosis and epidemiological studies of parasitic infections. *International Journal for Parasitology* **27,** 1135-1145.

Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy*. W.H. Freeman and Co., San Francisco.

Soares, L. R., Sercarz, E. E. and Miller, A. (1994) Vaccination of the *Leishmania major* susceptible BALB/c mouse. I. The precise selection of peptide determinant influences CD4+ T cell subset expression. *International Immunology* **6,** 785-794.

Sokal, R. R. and Sneath, P. H. A. (1963) *Principles of numerical taxonomy*. Freeman W.W. 359 pp.

Soto, M., Requena, J. M., Moreira, D. and Alonso, C. (1995) Assignment of genes to

Leishmania infantum chromosomes: karyotype and ploidy. *FEMS Microbiology Letters* **129,** 27-32.

Spencer, D. F., Collings, J. C., Schnare, M. N. and Gray, M. W. (1987) Multiple spacer sequences in the nuclear large subunit ribosomal RNA gene of *Crithidia fasciculata. The EMBO Journal* **6,** 1063-1071.

Steel, M. and Penny, D. (2000) Parsimony, likelihood, and the role of models in molecular phylogenetics. *Molecular Biology and Evolution* **17,** 839-850.

Steinkraus, H. B., Greer, J. M., Stephenson, D. C. and Langer, P. J. (1993) Sequence heterogeneity and polymorphic gene arrangements of the *Leishmania guyanensis* gp63 genes. *Molecular and Biochemical Parasitology* **62,** 173-185.

Stenger, S., Donhauser, N., Thuring, H., Rollinghoff, M. and Bogdan, C. (1996) Reactivation of latent leishmaniasis by inhibition of inducible nitric oxide synthase. *Journal of Experimental Medicine* **183,** 1501-1514.

Stevens, J. R. and Gibson, W. (1999) The molecular evolution of Trypanosomes. *Parasitology Today* **15,** 432-437.

Stothard, J. R., Frame, I. A., Carrasco, H. J. and Miles, M. A. (1998) On the molecular taxonomy of *Trypanosoma cruzi* using riboprinting. *Parasitology* **117,** 243-7.

Stothard, J. R., Frame, I. A. and Miles, M. A. (1997) Use of polymerase chain reaction-based single strand conformational polymorphism and denaturing gradient gel electrophoresis methods for detection of sequence variation of ribosomal DNA of *Trypanosoma cruzi*. *International Journal for Parasitology* **27,** 339-343.

Stothard, J. R., Hughes, S. and Rollinson, D. (1996) Variation within the internal transcribed spacer (ITS) of ribosomal DNA genes of intermediate snail hosts within the genus *Bulinus* (Gastropoda: Planorbidae). *Acta Tropica* **61,** 19-29.

Streit, J. A., Donelson, J. E., Agey, M. W. and Wilson, M. E. (1996) Developmental changes in the expression of *Leishmania chagasi* gp63 and heat shock protein in a human macrophage cell line. *Infection and Immunity* **64,** 1810-1818.

Thakur, C. P. (1984) Epidemiological, clinical and therapeutic features of Bihar kala-azar (including post kala-azar dermal leishmaniasis). *Transactions of the Royal Society of Tropical Medicine and Hygiene* **78,** 391-398.

Thakur, C. P. and Kumar, K. (1992) Post kala-azar dermal leishmaniasis: a neglected aspect of kala-azar control programmes. *Annals of Tropical Medicine and Parasitology* **86,** 355-359.

Thomaz Soccol, V., Lanotte, G., Rioux, J. A., Pratlong, F., Martini Dumas, A. and Serres, E. (1993) Monophyletic origin of the genus *Leishmania* Ross, 1903. *Annales de Parasitologie Humaine et Comparee* **68,** 107-108.

Tibayrenc, M. (1998) Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. *International Journal for Parasitology* **28,** 85-104.

Tibayrenc, M. and Ayala, F. J. (1991) Towards a population genetics of microorganisms:

the clonal theory of parasitic protozoa. *Parasitology Today* **7,** 228-232.

Tibayrenc, M., Kjellberg, F. and Ayala, F. J. (1990) A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas* and *Trypanosoma* and their medical and taxonomical consequences. *Proceedings of the National Academy of Sciences USA* **87,** 2414-2418.

Travi, B. L., Jaramillo, C., Montoya, J., Segura, I., Zea, A., Goncalves, A. and Velez, I. D. (1994) *Didelphis marsupialis*, an important reservoir of *Trypanosoma (Schizotrypanum) cruzi* and *Leishmania (Leishmania) chagasi* in Colombia. *American Journal of Tropical Medicine and Hygiene* **50,** 557-565.

Travi, B. L., Osorio, Y., Becerra, M. T. and Adler, G. H. (1998) Dynamics of *Leishmania chagasi* infection in small mammals of the undisturbed and degraded tropical dry forests of northern Colombia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92,** 275-278.

Tyagi, S., Bratu, D. P. and Kramer, F. R. (1998) Multicolor molecular beacons for allele discrimination. *Nature Biotechnology* **16,** 49-53.

Tyagi, S. and Kramer, F. R. (1996) Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology* **14,** 303-308.

van Eys, G. J., Guizani, I., Ligthart, G. S. and Dellagi, K. (1991) A nuclear DNA probe for the identification of strains within the *Leishmania donovani* complex. *Experimental Parasitology* **72,** 459-463.

van Eys, G. J., Schoone, G. J., Ligthart, G. S., Alvar, J., Evans, D. A. and Terpstra, W. J. (1989) Identification of 'Old World' *Leishmania* by DNA recombinant probes. *Molecular and Biochemical Parasitology* **34,** 53-62.

Vercesi, A. E., Moreno, S. N. and Docampo, R. (1994) Ca2+/H+ exchange in acidic vacuoles of *Trypanosoma brucei*. *Biochemical Journal* **304,** 227-233.

Vexenat, J. A., Fonseca de Castro, J. A., Cavalcante, R., Da Silva, M. R., Batista, W. H., Campos, J. H., Pereira, F. C., Tavares, J. P. and Miles, M. A. (1993) Preliminary observations on the diagnosis and transmissibility of canine visceral leishmaniasis in Teresina, n.e. Brazil. *Archives de l'Institut Pasteur de Tunis* **70,** 467-472.

Vicente, A. D. (1990) Kala-azar infatil. Dados epidemiológicos e incidência do kala-akar no internamento de pediatria do Hospital Distrital de Faro. *Revista Portuguesa de Doenças Infecciosas* **13,** 173-182.

Vickerman, K. (1974) Flagellate ultrastructure. In *Trypanosomiasis and leishmaniasis with special reference to Chagas disease.* Caracas, Venezuela. Elsevier.

Vickerman, K. (1994) The evolutionary expansion of the trypanosomatid flagellates. *International Journal for Parasitology* **24,** 1317-1331.

Victoir, K., Banuls, A. L., Arevalo, J., Llanos Cuentas, A., Hamers, R., Noel, S., De Doncker, S., Le Ray, D., Tibayrenc, M. and Dujardin, J. C. (1998) The gp63 gene locus, a target for genetic characterization of *Leishmania* belonging to subgenus *Viannia*. *Parasitology* **117,** 1-13.

Victoir, K., Dujardin, J. C., de Doncker, S., Barker, D. C., Arevalo, J., Hamers, R. and Le Ray, D. (1995) Plasticity of gp63 gene organization in *Leishmania (Viannia) braziliensis* and *Leishmania (Viannia) peruviana*. *Parasitology* **111,** 265-273.

Vila, C., Savolainen, P., Maldonado, J. E., Amorim, I. R., Rice, J. E., Honeycutt, R. L., Crandall, K. A., Lundeberg, J. and Wayne, R. K. (1997) Multiple and ancient origins of the domestic dog. *Science* **276,** 1687-1689.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and et al. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23,** 4407-14.

Voth, B. R., Kelly, B. L., Joshi, P. B., Ivens, A. C. and McMaster, W. R. (1998) Differentially expressed *Leishmania major* gp63 genes encode cell surface leishmanolysin with distinct signals for glycosylphosphatidylinositol attachment. *Molecular and Biochemical Parasitology* **93,** 31-41.

Wayne, R. K. (1993) Molecular evolution of the dog family. *Trends in Genetics* **9,** 218-224.

Walters, L. L. (1993) *Leishmania* differentiation in natural and unnatural sand fly hosts. *Journal of Eukaryotic Microbiology* **40,** 196-206.

Warburg, A., Saraiva, E., Lanzaro, G. C., Titus, R. G. and Neva, F. (1994) Saliva of *Lutzomyia longipalpis* sibling species differs in its composition and capacity to enhance leishmaniasis. *Philosophical Transactions of the Royal Society of London B Biological Sciences* **345,** 223-230.

Webb, J. R., Button, L. L. and McMaster, W. R. (1991) Heterogeneity of the genes encoding the major surface glycoprotein of *Leishmania donovani*. *Molecular and Biochemical Parasitology* **48,** 173-84.

Weir, B. S. (1996) *Genetic data analysis II. Methods for discrete population genetic data.* Sinauer Associates, Inc. Publishers, Sunderland. 445 pp.

Weiss, J. B. (1995) DNA probes and PCR for diagnosis of parasitic infections. *Clinical Microbiology Reviews* **8,** 113-130.

Welsh, J. and McClelland, M. (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* **18,** 7213-7218.

Wheeler, W. C. and Honeycutt, R. L. (1988) Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. *Molecular Biology and Evolution* **5,** 90-6.

WHO (1994) Leishmaniasis in the Americas. *Epidemiological Bulletin*, **15**, 8-13.

WHO, Expert Commitee (1990) Control of the leishmaniases. *WHO Technical Report Series* **793,** 1-158.

Widmer, G. and Dooley, S. (1995) Phylogenetic analysis of *Leishmania* RNA virus and *Leishmania* suggests ancient virus-parasite association. *Nucleic Acids Research* **23,** 2300-2304.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic*

Acids Research **18,** 6531-6535.

Wilson, I. B., O'Donnell, N., Allen, S., Mehlert, A. and Ferguson, M. A. (1999) Typing of *Leishmania* lipophosphoglycans by electrospray mass spectrometry. *Molecular and Biochemical Parasitology* **100,** 207-215.

Wilson, M. E. and Streit, J. A. (1996) Visceral leishmaniasis. *Gastroenterology Clinics of North America* **25,** 535-551.

Wilson, M. R. (1998) Apoptotic signal transduction: emerging pathways. *Biochemistry and Cell Biology* **76,** 573-582.

Wilson, S. M. (1995) DNA-based methods in the detection of *Leishmania* parasites: field applications and practicalities. *Annals of Tropical Medicine and Parasitology* **89 S1,** 95-100.

Wincker, P., Ravel, C., Blaineau, C., Pages, M., Jauffret, Y., Dedet, J. P. and Bastien, P. (1996) The *Leishmania* genome comprises 36 chromosomes conserved across widely divergent human pathogenic species. *Nucleic Acids Research* **24,** 1688-1694.

Wincker, P., Ravel, C., Britto, C., Dubessay, P., Bastien, P., Pages, M. and Blaineau, C. (1997) A direct method for the chromosomal assignment of DNA markers in *Leishmania*. *Gene* **194,** 77-80.

Wright, A. D., Li, S., Feng, S., Martin, D. S. and Lynn, D. H. (1999) Phylogenetic position of the kinetoplastids, *Cryptobia bullocki*, *Cryptobia catostomi*, and *Cryptobia salmositica* and monophyly of the genus *Trypanosoma* inferred from small subunit ribosomal RNA sequences. *Molecular and Biochemical Parasitology* **99,** 69-76.

Yang, D., Rogers, M. V., Brett, S. J. and Liew, F. Y. (1993) Immunological analysis of the zinc-binding peptides of surface metalloproteinase (gp63) of *Leishmania major*. *Immunology* **78,** 582-585.

Zijlstra, E. E., El Hassan, A. M., Ismael, A. and Ghalib, H. W. (1994) Endemic kala-azar in eastern Sudan: a longitudinal study on the incidence of clinical and subclinical infection and post kala-azar dermal leishmaniasis. *American Journal of Tropical Medicine and Hygiene* **51,** 826-836.

10. Annexes

10.1. Annex 1 - PCR

10.1.1. Reaction Mixtures

All PCR mixtures were prepared fresh from all reagents and in a master solution for the total amount of reactions in each series. Water was always milliQ (mQ), filtered sterilized through 0.2 μm syringe filters and autoclaved. The 10x NH_4 buffer and $MgCl_2$ solution were supplied with *Taq* DNA Polymerase (Bioline) and were used in all PCR reactions, except for ITS (670mM Tris HCl, pH8.8) and mini-exon (0.2M Tris-HCl, 250mM KCl, pH8.3) amplification. The reactions were prepared for a default total volume of 20μl. All PCR reactions had 1.25ng/ul of genomic DNA unless stated otherwise.

Table XXXIII - Composition of optimised PCR and RAPD reactions. Only one primer is used in each RAPD.

REAGENTS		$MgCl_2$ (mM)	dNTPs (mM)	Primer (μM)	DMSO (%)	BSA (g/l)	Tween20 (% V/V)	*Taq* DNA polymerase (U/100 μl)
RAPD		2.5	0.2	0.8	-	-	-	5
ITSa		2.5	0.2	0.5 + 0.5	-	-	-	2.5
ITSb		3.0	0.2	0.03 + 0.03	-	-	-	5
	ITG/L	2.5	0.2	0.5 + 0.5	5	-	-	5
	ITG/CS	2.5	0.2	0.5 + 0.5	10	-	-	5
gp63	**A**	2.0	0.2	0.5 + 0.5	-	-	-	5
	B	4.0	0.2	0.5 + 0.5	5	-	-	5
	C	4.0	0.2	0.5 + 0.5	5	-	-	5
	mspC3	2.0	0.2	0.5 + 0.5	5	-	-	5
Mini-exon a		1.5	0.225	1 + 1	10	0.1	0.5	12.5
Mini-exon b		1.5	0.225	0.25 + 0.25	10	0.1	0.5	12.5

NOTES: ITSb was optimised to amplify *Leishmania* DNA from blotted human biopsies. Mini-exon b was used to amplify 250ng of DNA in a 100 μl reaction.

10.1.2. Amplification cycles

Every PCR was carried out in a Hybaid Omnigene thermal reactor.

Table XXXIV - Thermal cycling conditions for RAPD, ITS and mini-exon.

STEP	RAPD			ITS			Mini-exon		
	T (°C)	time (min)	number of cycles	T (°C)	time (sec)	number of cycles	T (°C)	time (min)	number of cycles
denaturation	-	-	-	-	-	-	95	5	1
denaturation	94	5		95	15		95	1	5
annealing	37	1	1	58	30	37	45	0.5	
extension	72	1		72	90		65	1	
denaturation	94	1		-	-		95	1	35
annealing	37	1	37	-	-	-	50	0.5	
extension	72	1		-	-		72	1	
extension	72	1	1	-	-	-	72	10	1

T - temperature

Table XXXV - Thermal cycling conditions for gp63 PCR amplifications.

STEP	gp63 ITG/L			*gp63 ITG/CS*			*mspC3* A			*mspC3* B &C			*mspC3*		
	T (°C)	time (min)	number of cycles	T (°C)	time (min)	number of cycles	T (°C)	time (min)	number of cycles	T (°C)	time (min)	number of cycles	T (°C)	time (min)	number of cycles
denaturation	95	1	1	94	5	1	94	5	1	94	5	1	94	5	1
denaturation	95	1		94	1		94	1		94	1		94	1	
annealing	70	1	35	65	1	30	55	1	30	65	1	5	55	1	30
extension	72	1.5		72	2		72	1		72	1		72	1	
denaturation	-	-		-	-		-	-		94	1		-	-	
annealing	-	-	-	-	-	-	-	-	-	60	1	25	-	-	-
extension	-	-		-	-		-	-		72	1		-	-	
extension	72	10	1	72	10	1	72	10	1	72	10	1	72	10	1

T - temperature

10.2. Annex 2 - Data tables

Table XXXVI - Binary and multistate ('sequence') coded data from isoenzyme profiles for analysed zymodemes.

	Zymodeme (ILM)														
Enzyme	**1**	**3**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**16**	**17**	**18**
ALAT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
ASAT	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	1	1	1	1	0	1	0	1	0	1	0	1	1	0	0
	0	0	0	1	1	0	1	0	1	1	0	0	0	1	0
GPI	1	1	1	1	1	1	1	0	1	0	0	1	0	0	0
	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0
	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MPI	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
NH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	1	1	1	1	1	0	0	1	1	1	0	0	0	0	0
	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
PEPD	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
6PGD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
MDH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	1	1	1	1	1	0	1	1	0	0
ALAT	a	a	a	a	a	a	a	a	a	a	t	c	t	g	g
ASAT	a	a	a	at	t	a	t	a	t	at	t	a	a	t	g
GPI	a	a	a	a	a	a	a	t	a	a	g	a	t	g	g
MDH	a	a	t	t	t	a	a	a	a	a	g	a	a	g	g
MPI	a	a	a	a	a	a	a	t	a	a	g	a	a	g	g
NH	a	a	a	a	a	t	t	a	a	a	g	t	c	g	g
PEPD	a	a	a	a	a	a	a	t	a	a	g	a	a	g	g
6PGD	a	t	t	t	t	t	t	t	t	t	at	t	t	g	g

Table XXXVII - Data matrices for RAPD of *L. infantum* and *L. chagasi*. 280 bands scored for 33 strains.

Primer A2

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D2	D3	D4
	1	1	0	0	0	1	0	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	1	1	0	1	0	1	0	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	5	5	2	7	7	8	7	8	8	8	8	8	8	7	5	7	8	7	7	8	7	6	8	8	7	8	7	8	7	9	7	6	8

In order, from left to right: outgroup: *L. major* (M): 5-ASKH; *L. aethiopica* (A): L96; *L. tropica* (T): K27. *L. infantum* (I-I17): IPT-1; L82; Pharoah; Lombardi; LEM75; IMPT104; IMT150; IMT152; IMT89; IMT124; IMT108; IMT193; Strain A; L53. *L. chagasi* (C-C12): PP75; WR285; CO910; M9702; M8270; M12727; M12734; M12337; M7633; M12085; M12084. *L. donovani* (D- D4): DD8; HU3; MRC(L)3; MRC74; Patna1.

Primer A4

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	0	1	0	1	1	0	0	1	1	1	1	1	1	0	1	0	0	1
	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	0	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1
	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0
	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0
	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
	0	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1
	1	1	1	1	1	1	1	1	0	0	0	0	0	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0
total	7	3	9	12	9	9	6	11	9	10	9	10	9	11	10	9	10	10	12	11	10	6	10	10	9	9	9	9	10	12	11	11	

Primer A5

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	0	1	1	1	1	1	0	1	1	0	0	0	1	1	1	0	1	0	1	1	0	1	0	0	0	0	0
	0	0	0	1	1	0	0	0	0	0	0	1	0	1	1	0	0	0	1	1	0	0	0	0	0	0	1	1	1	0	1	0	1
	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
total	1	0	4	6	5	6	4	4	5	6	5	7	3	7	7	3	3	3	5	7	4	2	6	4	6	5	5	7	5	2	6	5	5

Primer A6

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	10	13	17	11	11	11	10	11	11	11	11	11	11	11	11	11	10	11	10	11	11	10	11	11	11	11	11	11	12		10	13	12

Primer D3

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
total	7	4	8	6	7	6	7	6	6	6	6	6	6	6	6	6	6	6	7	6	7	7	6	6	6	7	6	6	8	6	9	6	8

Primer D8

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	0	1	0	0	1	1	0	1	0	0	0	1
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	1	1	1
	0	0	0	1	1	1	1	1	1	1	0	0	0	1	0	0	1	1	1	1	1	1	1	0	0	1	0	0	0	1	0	0	0
	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	1	1	0	0	1	0	0	1	1	0	0	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	13	14	13	14	13	15	12	12	13	16	14	13	13	15	14	12	14	15	13	13	15	15	16	14	13	16	14	13	13	11	10	12	13

Primer D10

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	0	1	1	1	0	0	0	0	1
	1	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	1	1	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1
	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	5	7	7	10	9	12	6	10	11	11	11	12	11	12	11	12	12	12	10	10	10	9	12	12	11	13	12	12	8	8	10	11	

Primer H1

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	0	0	0	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
total	7	6	9	11	9	13	9	10	13	13	13	13	13	13	13	13	12	10	11	10	11	9	11	11	11	11	12	11	13	12	12	12	

Primer H4

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	1	1	1	1	1	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	14	10	11	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	11	12	13	13	13	14	11	13	13	13	13	12	11

Primer L2

Strain	M	A	T	I	I1	I3	I4	I5	I6	I8	I9	I10	I11	I12	I13	I16	I17	C	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	D	D1	D 2	D 3	D4
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	0	1	0	1	1	1	0
	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
total	13	12	19	13	13	13	13	13	13	13	13	13	13	12	13	13	13	12	13	13	12	13	13	13	13	13	12	13	13	15	14	14	12

Table XXXVIII - Data matrices for RAPD of Portuguese *L. infantum*. 280 bands scored for 33 strains.

Primer A2

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	1	0	0	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0
	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	1	1	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
total	4	7	6	5	7	6	6	7	7	6	5	6	8	2	6	6	9	8	8	8	5	5	6	6

In order, from left to right: Reference strains: M - 5-ASKH; A - L96; D - DD8; C - PP75; I - IPT-1. Portuguese strains: 1- L82; 2 - IMT104; 3 - IMT150; 4 - IMT152; 5 - IMT89; 6 - IMT124; 7 - IMT108; 8 - IMT193; 9 - IMT160; 10 - IMT162; 11 - IMT161; 12- IMT169; 13 - IMT170; 14 - IMT171; 15 - IMT172; 16 - IMT177; 17 - IMT191; 18 - IMT195.

Primer A4

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	6	7	8	10	10	10	9	8	9	8	9	9	9	10	10	10	11	11	12	12	12	11	11	11

Primer A5

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	0	0	0	0	0	0	1	1	0	0	1	1	1	0	1	1	1	1	1	1
	0	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0
	0	0	1	0	1	1	0	0	0	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1
	0	0	0	0	1	0	0	1	0	1	0	1	1	1	1	1	0	0	0	1	1	1	0	1
	0	0	1	0	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
total	2	4	6	4	5	3	3	6	5	6	4	8	7	6	7	8	7	6	6	8	6	8	6	7

Primer A6

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	1	1	0	1	1
	1	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
total	12	9	8	8	7	7	9	10	10	10	10	10	10	8	10	8	9	9	8	9	9	8	10	10

Primer D3

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	0	1	1	1	0	1	1	1	0	0	1	1	0	0	1	0	0	1	1	0
	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	1	1	0	1	1	0	0	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	7	6	7	10	8	6	7	8	8	8	8	8	8	8	7	8	8	8	8	8	8	8	8	8

Primer D8

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	0	0	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0	0	1	0
	0	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0	0	1	1
	0	0	1	1	0	0	0	0	1	0	0	0	1	0	1	1	1	0	0	1	0	0	0	0
	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
total	5	0	7	6	5	5	9	10	11	10	9	10	13	8	11	11	11	8	8	11	8	8	10	8

Primer D10

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	0	0	0	0	0
	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	8	9	13	14	13	13	14	14	14	14	13	14	14	14	14	14	13	13	14	13	13	13	13	13

Primer H1

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	0	0	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	0	0	0	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	1	0	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	0	0	1	0	1	1	1	1	0	-	1	1	0	0	0	0	0	0	0	1
	0	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	0
	1	0	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	0	0	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	1	0	1	1	0	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	0	0	1	1
	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
total	9	11	16	15	12	13	14	14	15	15	15	15	14	-	15	15	14	14	14	14	13	13	14	14

Primer L2

strain	M	A	D	C	I	I1	I6	I8	I9	I10	I11	I12	I13	I18	I20	I21	I22	I23	I24	I25	I26	I27	I29	I30
	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	0	0	1	0	1	1	0	0	0	1	1	0	1	1	1	1	0	0	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
total	3	3	3	4	3	3	4	5	5	4	4	4	5	5	4	5	5	5	5	4	4	5	5	5

Table XXXIX - Data matrices for RAPD amplification within the *L. donovani* complex.

Primer A2.

strain	bands																total
M	1	1	0	0	0	0	0	1	0	1	1	0	0	0	1	0	6
A	1	0	0	1	0	0	1	1	1	0	0	0	0	1	0	0	6
C	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
I	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
D	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D1	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D2	0	0	0	0	0	1	0	1	1	0	0	0	1	0	0	1	5
D3	0	0	0	0	0	1	0	1	1	0	0	0	1	0	0	1	5
D4	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D5	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D6	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
D7	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D9	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	4
D10	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	4
D11	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	3
D12	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	3
D13	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	3
D14	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	4
D15	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	4
D16	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
D17	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	3
D18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D19	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D20	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	4
D21	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	4
D22	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	4
D23	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	4
D24	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D25	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D26	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D27	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D28	0	0	0	0	0	1	0	0	1	0	1	0	1	0	0	1	5
D29	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D30	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
D31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D32	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
D33	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
D34	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
D35	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
I31	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
I4	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
I16	0	0	0	1	0	1	0	1	1	0	1	0	1	0	0	1	7
I17	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
I25	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6
I26	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	6

Primer A4.

strain	bands																								total
M	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	1	6
A	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
C	0	1	1	0	0	1	1	0	0	0	0	0	1	1	0	1	0	0	1	1	1	0	0	0	10
I	0	1	1	0	0	1	1	0	0	0	0	1	1	1	0	1	0	0	0	1	1	0	0	0	10
D	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	7
D1	0	1	1	0	0	0	1	0	1	0	0	0	1	1	0	1	0	0	1	1	0	0	0	0	9
D2	0	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	7
D3	0	1	1	0	0	0	1	1	0	1	0	0	0	0	0	1	0	1	1	0	0	1	0	0	9
D4	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	5
D5	0	1	1	0	0	0	1	0	0	1	0	0	1	1	0	0	0	1	0	1	0	1	0	0	9
D6	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	0	7
D7	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	0	7
D8	0	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	1	0	0	0	0	0	7
D9	0	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	1	0	0	0	0	0	7
D10	0	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	1	1	0	0	0	0	0	8
D11	0	1	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	5
D12	0	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	1	1	0	0	0	0	0	8
D13	0	1	1	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	8
D14	0	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	1	1	0	0	0	0	0	8
D15	0	1	1	0	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	6
D16	0	1	1	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	7
D17	0	1	1	0	0	0	1	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	6
D18	0	1	1	0	0	0	1	0	1	0	1	0	0	1	0	0	0	1	1	0	0	1	0	0	9
D19	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	6
D20	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	7
D21	0	1	1	0	0	0	1	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	8
D22	0	1	1	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	8
D23	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	5
D24	0	1	1	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	1	1	0	1	0	0	8
D25	0	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	1	0	0	8
D26	0	1	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	7
D27	0	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	7
D28	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	7
D29	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	5
D30	0	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	1	0	0	8
D31	0	1	1	0	0	0	1	1	0	1	0	0	0	1	0	1	1	1	0	0	0	1	0	0	10
D32	0	1	1	0	0	0	1	0	1	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	8
D33	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	6
D34	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	6
D35	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	6
I31	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	6
I4	0	1	1	0	0	1	0	0	0	0	1	0	1	1	0	1	0	0	1	1	0	0	0	0	9
I16	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0	0	0	8
I17	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0	0	0	8
I25	0	1	1	0	0	1	0	0	0	0	1	0	1	1	0	1	0	1	0	1	0	0	0	0	9
I26	0	1	1	0	0	1	0	0	0	0	1	0	1	1	0	1	0	1	0	1	0	0	0	0	9

Primer A5.

strain	bands																	total
M	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	2
A	0	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	4
C	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	4
I	0	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	8
D	0	1	0	0	0	0	0	0	1	1	0	1	1	1	0	0	1	7
D1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	6
D2	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	5
D3	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	4
D4	0	1	0	0	0	0	0	0	1	1	0	1	1	1	0	0	1	7
D5	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	4
D6	0	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	1	5
D7	0	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	1	5
D8	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	4
D9	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	5
D10	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	5
D11	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	4
D12	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	5
D13	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	5
D14	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	5
D15	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	5
D16	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	4
D17	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2
D18	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2
D19	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	6
D20	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	4
D21	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	3
D22	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	3
D23	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
D24	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
D25	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	4
D26	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	3
D27	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	4
D28	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	4
D29	0	1	0	0	0	0	0	0	0	1	0	1	1	1	0	0	1	6
D30	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	4
D31	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	3
D32	0	1	0	0	0	0	0	0	0	1	0	1	1	1	0	0	1	6
D33	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	1	5
D34	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	3
D35	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	3
I31	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	4
I4	0	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	6
I16	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	4
I17	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	4
I25	0	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	7
I26	0	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	7

Primer A6

strain	bands													total
M	0	1	1	0	0	1	1	0	0	0	0	0	0	4
A	0	1	0	0	1	0	0	1	0	0	0	0	0	3
C	0	0	0	0	0	1	0	1	0	0	1	1	0	4
I	0	0	0	0	0	1	0	1	0	0	1	1	0	4
D	0	0	0	0	0	1	0	1	1	1	1	1	0	6
D1	0	0	0	0	0	1	0	1	1	1	0	1	0	5
D2	0	0	0	0	0	1	0	1	1	1	1	0	0	5
D3	0	0	0	0	0	1	0	1	1	1	1	0	0	5
D4	0	0	0	0	0	1	0	1	1	1	1	1	0	6
D5	0	0	0	0	0	1	0	1	1	1	0	1	0	5
D6	0	0	0	0	0	1	0	1	1	1	1	0	0	5
D7	0	0	0	0	0	1	0	1	1	1	1	0	0	5
D8	0	0	0	0	0	1	0	1	0	1	1	0	0	4
D9	-	-	-	-	-	-	-	-	-	-	-	-	-	0
D10	0	0	0	0	0	1	0	1	0	1	1	0	0	4
D11	0	0	0	0	0	1	0	1	0	1	1	0	0	4
D12	0	0	0	0	0	1	0	1	0	0	0	0	0	2
D13	0	0	0	0	0	1	0	1	0	0	0	0	0	2
D14	0	0	0	0	0	1	0	1	0	1	1	0	0	4
D15	0	0	0	0	0	1	0	1	0	1	1	0	0	4
D16	0	0	0	0	0	1	0	1	0	0	1	0	0	3
D17	0	0	0	0	0	1	0	1	1	1	0	0	0	4
D18	0	0	0	0	0	1	0	1	1	1	0	0	0	4
D19	0	0	0	0	0	1	0	1	0	0	0	0	0	2
D20	0	0	0	0	0	1	0	1	0	1	1	0	0	4
D21	0	0	0	0	0	1	0	1	0	0	0	0	0	2
D22	0	0	0	0	0	1	0	1	0	0	0	0	0	2
D23	1	1	0	1	0	1	0	1	0	1	0	0	1	7
D24	0	0	0	0	0	1	0	1	1	1	0	1	0	5
D25	0	0	0	0	0	1	0	1	1	1	1	1	0	6
D26	0	0	0	0	0	1	0	1	1	1	1	1	0	6
D27	0	0	0	0	0	1	0	1	1	1	1	0	0	5
D28	0	0	0	0	0	1	0	1	1	1	0	0	0	4
D29	0	0	0	0	0	1	0	1	0	0	1	0	0	3
D30	0	0	0	0	0	1	0	1	0	0	1	0	0	3
D31	0	0	0	0	0	1	0	1	1	0	1	0	0	4
D32	0	0	0	0	0	1	0	1	1	0	1	1	0	5
D33	0	0	0	0	0	1	0	1	1	0	0	0	0	3
D34	0	0	0	0	0	1	0	1	1	0	0	1	0	4
D35	0	0	0	0	0	1	0	1	1	0	0	1	0	4
I31	0	0	0	0	0	1	0	1	0	0	1	0	0	3
I4	0	0	0	0	0	1	0	1	1	0	1	0	0	4
I16	0	0	0	0	0	1	0	1	1	0	1	0	0	4
I17	0	0	0	0	0	1	0	1	1	0	1	0	0	4
I25	0	0	0	0	0	1	0	1	0	0	1	0	0	3
I26	0	0	0	0	0	1	0	1	1	0	1	0	0	4

Primer D8

strain	bands														total
M	1	0	0	0	0	1	0	0	0	0	0	0	0	0	2
A	0	0	0	0	1	0	1	0	0	0	0	0	0	0	2
C	0	0	0	0	1	0	0	1	0	0	0	0	1	1	4
I	0	1	0	0	1	0	0	1	0	0	0	0	1	1	5
D	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
D1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	3
D2	0	0	1	1	1	0	0	1	1	0	0	0	0	0	5
D3	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D4	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D5	0	0	1	0	1	0	0	1	0	0	0	0	0	0	3
D6	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
D7	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D8	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D9	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D10	0	0	1	0	1	1	0	1	1	0	0	0	0	0	5
D11	0	0	1	0	1	1	0	1	1	0	0	0	0	0	5
D12	0	0	1	0	1	1	0	1	0	0	0	0	0	0	4
D13	0	0	1	0	1	1	0	1	1	0	0	0	0	0	5
D14	0	0	1	0	1	1	0	1	1	0	0	0	0	0	5
D15	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D16	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D17	0	0	1	0	1	0	0	1	1	0	1	0	1	0	6
D18	0	0	0	0	0	0	0	1	0	0	0	0	1	0	2
D19	0	0	1	0	1	0	0	1	0	0	0	0	0	0	3
D20	0	0	1	0	1	0	0	1	0	0	1	0	0	0	4
D21	0	0	1	0	1	0	0	1	0	0	1	0	0	0	4
D22	0	0	1	0	1	0	0	1	0	0	0	0	0	0	3
D23	0	0	1	0	1	1	0	0	1	0	0	0	0	0	4
D24	0	0	1	1	0	0	0	1	0	0	0	0	0	0	3
D25	0	0	1	0	1	0	0	1	0	1	0	0	0	0	4
D26	0	0	1	0	1	0	0	1	0	0	0	1	0	0	4
D27	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D28	0	0	1	0	1	0	0	1	1	0	0	0	1	0	5
D29	0	0	1	0	1	1	0	1	1	0	0	0	0	0	5
D30	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D31	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D32	0	0	1	0	1	0	0	1	1	0	0	0	0	0	4
D33	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
D34	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
D35	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
I31	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
I4	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
I16	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
I17	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
I25	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4
I26	0	0	1	0	1	0	0	1	0	0	0	0	1	0	4

Primer D10

strain	bands												total
M	0	0	0	0	0	1	0	0	0	0	1	0	2
A	1	0	0	0	0	0	1	0	1	0	0	0	3
C	0	0	0	1	1	0	0	0	0	1	0	1	4
I	0	1	1	1	1	0	0	0	0	1	0	1	6
D	0	0	0	1	1	0	0	0	0	1	0	1	4
D1	0	1	0	1	1	0	0	0	0	1	0	1	5
D2	0	1	0	0	1	0	0	0	0	1	0	1	4
D3	0	1	0	0	1	0	0	0	0	1	0	1	4
D4	0	0	0	1	1	0	0	0	0	1	0	1	4
D5	0	1	0	1	1	0	0	0	0	1	0	1	5
D6	0	0	0	1	1	0	0	0	0	1	0	1	4
D7	0	0	0	1	1	0	0	0	0	1	0	1	4
D8	0	0	0	1	1	0	0	0	0	1	0	0	3
D9	0	0	0	0	0	0	0	0	0	0	0	0	0
D10	0	0	0	1	1	0	0	0	1	1	0	1	5
D11	0	0	0	0	0	0	0	0	1	0	0	0	1
D12	0	1	0	1	1	0	0	0	0	1	0	1	5
D13	0	1	0	1	1	0	0	0	0	1	0	1	5
D14	0	0	0	1	1	0	0	0	0	1	0	1	4
D15	0	0	0	0	1	0	0	0	0	1	0	1	3
D16	0	1	0	0	1	0	0	0	0	1	0	1	4
D17	0	1	0	0	1	0	0	0	1	1	0	0	4
D18	0	0	0	1	1	0	0	0	0	1	0	1	4
D19	0	1	0	1	1	0	0	0	0	1	0	1	5
D20	0	1	0	1	1	0	0	0	0	1	0	1	5
D21	0	1	0	0	1	0	0	0	0	1	0	1	4
D22	0	1	0	1	1	0	0	0	0	1	0	1	5
D23	0	1	0	0	0	0	1	0	0	1	0	0	3
D24	0	1	0	1	1	0	0	0	0	1	0	1	5
D25	0	1	0	0	1	0	0	0	0	1	0	1	4
D26	0	1	0	1	1	0	0	0	0	1	0	1	5
D27	0	1	0	0	1	0	0	0	0	1	0	1	4
D28	0	1	0	1	1	0	0	0	0	1	0	0	4
D29	0	1	0	1	1	0	0	0	0	1	0	1	5
D30	0	1	0	0	1	0	0	0	0	1	0	1	4
D31	0	1	0	0	1	0	0	0	0	1	0	1	4
D32	0	1	0	1	1	0	0	0	0	1	0	1	5
D33	0	1	0	1	1	0	0	0	0	1	0	1	5
D34	0	1	0	1	1	0	0	0	0	1	0	1	5
D35	0	1	0	1	1	0	0	0	0	1	0	1	5
I31	0	1	0	1	1	0	0	0	0	1	0	1	5
I4	0	0	0	1	1	0	0	0	0	1	0	1	4
I16	0	0	0	1	1	0	0	0	0	1	0	1	4
I17	0	1	0	1	1	0	0	0	0	1	0	1	5
I25	0	1	0	1	1	0	0	0	0	1	0	1	5
I26	0	1	0	1	1	0	0	0	0	1	0	1	5

Primer H1

strain	bands													total
M	0	0	0	0	0	0	1	1	0	0	1	1	1	5
A	0	0	0	1	0	0	1	0	1	1	0	0	0	4
C	1	0	0	0	0	1	1	0	0	0	0	1	0	4
I	1	1	0	0	0	1	1	0	0	0	0	1	0	5
D	1	1	0	1	0	1	1	0	0	0	1	0	0	6
D1	1	0	0	0	0	1	1	0	0	0	1	1	0	5
D2	1	1	0	0	0	1	1	0	0	0	1	1	0	6
D3	1	1	0	0	0	0	1	0	0	0	1	1	0	5
D4	1	0	0	1	0	1	1	0	0	0	1	0	0	5
D5	1	1	0	0	0	1	1	0	0	0	1	1	0	6
D6	1	1	0	1	0	1	1	0	0	0	1	0	0	6
D7	1	0	0	1	0	1	1	0	0	0	1	0	0	5
D8	1	0	0	1	0	0	1	0	0	0	1	0	0	4
D9	0	0	0	0	1	0	1	0	0	0	1	1	0	4
D10	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D11	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D12	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D13	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D14	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D15	1	0	0	0	0	0	1	0	0	0	1	1	0	4
D16	1	0	0	0	0	1	1	0	0	0	1	1	0	5
D17	0	0	0	0	0	0	0	0	0	0	1	1	0	2
D18	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D19	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D20	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D21	1	1	0	0	0	1	1	0	0	0	1	1	0	6
D22	1	0	0	0	1	1	1	0	0	0	1	1	0	6
D23	0	0	1	1	0	0	0	0	1	0	0	0	0	3
D24	1	0	0	0	0	1	1	0	0	0	1	1	0	5
D25	1	1	0	0	0	1	1	0	0	0	1	1	0	6
D26	1	0	0	0	0	0	1	0	0	0	1	0	0	3
D27	1	1	0	0	0	0	1	0	0	0	1	1	0	5
D28	1	1	0	0	0	1	1	0	0	0	1	1	0	6
D29	1	0	0	0	0	1	1	0	0	0	1	1	0	5
D30	1	1	0	0	0	0	1	0	0	0	1	1	0	5
D31	1	1	0	0	0	1	1	0	0	0	1	1	0	6
D32	1	0	0	0	1	0	1	0	0	0	1	1	0	5
D33	1	1	0	0	1	0	1	0	0	0	1	1	0	6
D34	1	1	0	0	1	0	1	0	0	0	1	1	0	6
D35	1	1	0	0	1	0	1	0	0	0	1	1	0	6
I31	1	1	0	0	0	1	1	0	0	0	0	1	0	5
I4	1	0	0	0	0	1	1	0	0	0	0	1	0	4
I16	1	1	0	0	0	1	1	0	0	0	0	1	0	5
I17	1	1	0	0	0	1	1	0	0	0	0	1	0	5
I25	1	1	0	0	0	1	1	0	0	0	0	1	0	5
I26	1	1	0	0	0	1	1	0	0	0	0	1	0	5

Primer H4

strain	bands													total
M	0	0	1	0	0	0	0	0	1	1	1	0	0	4
A	0	1	1	0	0	0	0	0	0	0	0	0	0	2
C	0	0	1	0	0	0	1	0	1	0	0	0	1	4
I	1	0	1	0	0	1	0	0	1	0	0	0	1	5
D	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D1	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D2	0	0	1	0	0	1	1	0	1	0	0	1	1	6
D3	0	0	1	0	0	1	0	0	1	0	0	1	1	5
D4	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D5	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D6	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D7	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D10	0	1	1	1	0	0	0	0	0	0	0	0	0	3
D11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D12	0	0	0	0	0	1	0	0	1	0	0	0	0	2
D13	0	0	1	0	0	1	0	0	1	0	0	0	0	3
D14	0	0	1	0	0	1	0	0	1	0	0	0	0	3
D15	0	1	1	1	0	0	0	0	1	0	0	1	0	5
D16	0	0	1	1	0	0	1	0	1	0	0	1	1	6
D17	0	0	1	0	0	0	0	0	1	0	0	0	0	2
D18	0	0	1	1	0	1	0	0	1	0	0	0	1	5
D19	0	0	1	0	0	1	0	0	1	0	0	0	1	4
D20	0	1	1	1	0	1	0	0	1	0	0	0	0	5
D21	0	0	1	0	0	1	1	0	1	0	0	1	0	5
D22	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D23	1	1	1	1	1	0	0	1	0	0	0	0	0	6
D24	0	0	1	0	0	1	1	0	1	0	0	0	1	5
D25	0	0	1	1	0	0	1	0	1	0	0	1	1	6
D26	0	0	1	1	0	0	1	0	1	0	0	1	1	6
D27	0	0	1	0	0	1	0	0	1	0	0	1	0	4
D28	0	0	1	1	0	1	1	0	1	0	0	0	1	6
D29	0	0	1	1	0	1	1	0	1	0	0	0	1	6
D30	0	0	1	1	0	1	1	0	1	0	0	1	1	7
D31	0	0	1	1	0	1	0	0	1	0	0	1	1	6
D32	0	0	1	1	0	1	1	0	1	0	0	0	1	6
D33	0	0	1	1	0	1	1	0	1	0	0	0	1	6
D34	0	0	1	1	0	0	1	0	1	0	0	0	1	5
D35	0	0	1	1	0	0	1	0	1	0	0	0	1	5
I31	0	0	1	1	0	0	1	0	1	0	0	0	1	5
I4	0	0	1	1	0	0	0	0	1	0	0	0	1	4
I16	0	0	1	1	0	0	1	0	1	0	0	0	1	5
I17	0	0	1	1	0	0	1	0	1	0	0	0	1	5
I25	0	0	1	1	0	0	1	0	1	0	0	0	1	5
I26	0	0	1	1	0	0	1	0	1	0	0	0	1	5

Primer L2

strain	bands										total
M	1	0	1	1	0	1	1	0	0	1	6
A	1	0	1	0	1	0	0	1	1	0	5
C	1	1	0	0	0	0	1	1	0	0	4
I	1	1	0	0	0	0	1	1	0	0	4
D	1	1	0	0	1	0	0	1	0	0	4
D1	1	1	0	0	0	1	0	1	0	0	4
D2	1	1	0	0	0	1	0	1	0	0	4
D3	1	1	0	0	0	1	0	1	0	0	4
D4	1	1	0	0	1	0	0	1	0	0	4
D5	1	1	0	0	1	0	0	1	0	0	4
D6	1	1	0	0	1	0	0	1	0	0	4
D7	1	1	0	0	1	0	0	1	0	0	4
D8	1	1	0	0	0	0	1	1	0	0	4
D9	1	1	0	0	0	1	1	1	0	0	5
D10	1	1	0	0	0	1	1	1	0	0	5
D11	1	1	0	0	0	1	1	1	0	0	5
D12	1	1	0	0	1	0	1	1	0	0	5
D13	1	1	0	0	0	1	1	1	0	0	5
D14	1	1	0	0	0	1	1	1	0	0	5
D15	1	1	0	0	0	1	1	1	0	0	5
D16	1	1	0	0	0	1	0	1	0	0	4
D17	1	1	0	0	0	1	0	1	0	0	4
D18	1	1	0	0	0	1	1	1	0	0	5
D19	1	1	0	0	0	1	0	1	0	0	4
D20	1	1	0	0	0	1	0	1	0	0	4
D21	1	1	0	0	0	1	0	1	0	0	4
D22	1	1	0	0	0	1	0	1	0	0	4
D23	0	0	0	0	0	0	0	0	0	0	0
D24	1	1	0	0	1	0	0	1	0	0	4
D25	1	1	0	0	0	1	0	1	0	0	4
D26	1	1	0	0	0	1	0	1	0	0	4
D27	1	1	0	0	0	1	0	1	0	0	4
D28	1	1	0	0	0	1	0	1	0	0	4
D29	1	1	0	0	0	1	0	1	0	0	4
D30	1	1	0	0	0	1	0	1	0	0	4
D31	1	1	0	0	0	1	0	1	0	0	4
D32	1	1	0	0	0	1	0	1	0	0	4
D33	1	1	0	0	0	1	0	1	0	0	4
D34	1	1	0	0	0	1	0	1	0	0	4
D35	1	1	0	0	0	1	0	1	0	0	4
I31	1	1	0	0	0	1	0	1	0	0	4
I4	1	1	0	0	0	1	0	1	0	0	4
I16	1	1	0	0	0	1	0	1	0	0	4
I17	1	1	0	0	0	1	0	1	0	0	4
I25	1	1	0	0	0	1	0	1	0	0	4
I26	1	1	0	0	0	1	0	1	0	0	4

Table XL - Total number of bands scored per each RAPD primer for *L. infantum* and *L. chagasi*.

	Strain																														
Primer	**M***	**A***	**T***	**I***	**I1**	**I3**	**I4**	**I5**	**I6**	**I8**	**I9**	**I10**	**I11**	**I12**	**I13**	**I16**	**I17**	**C***	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C11**	**C12**	**D***	**D1**	**D2**	**D3**	**D4**
A2	5	5	2	7	7	8	7	8	8	8	8	8	8	7	5	7	8	7	7	8	7	6	8	8	7	8	7	8	7	9	7	6	8
A4	7	3	9	12	9	9	6	11	9	10	9	10	9	11	10	9	10	10	12	11	10	6	10	10	9	9	9	9	10	12	11	11	12
A5	1	0	4	6	5	6	4	4	5	6	5	7	3	7	7	3	3	3	5	7	4	2	6	4	6	5	5	7	5	2	6	5	5
A6	10	13	17	11	11	11	10	11	11	11	11	11	11	11	11	11	10	11	10	11	11	10	11	11	11	11	11	11	12	10	13	12	13
D3	7	4	8	6	7	6	7	6	6	6	6	6	6	6	6	6	6	6	7	6	7	7	6	6	6	7	6	6	8	6	9	6	8
D8	13	14	13	14	13	15	12	12	13	16	14	13	13	15	14	12	14	15	13	13	15	15	16	14	13	16	14	13	13	11	10	12	13
D10	5	7	7	10	9	12	6	10	11	11	11	12	11	12	11	12	12	12	10	10	10	9	12	12	11	13	12	12	8	8	10	11	11
H1	7	6	9	11	9	13	9	10	13	13	13	13	13	13	13	13	12	10	11	10	11	9	11	11	11	11	12	11	13	12	12	12	13
H4	14	10	11	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	11	12	13	13	13	14	11	13	13	13	13	12	11
L2	13	12	19	13	13	13	13	13	13	13	13	13	13	12	13	13	13	12	13	13	12	13	13	13	13	13	12	13	13	15	14	14	12
total	82	74	99	103	96	106	87	98	102	107	103	106	100	108	103	99	101	89	101	102	98	89	106	102	100	107	99	103	102	98	105	101	106

A - *L. aethiopica*; C - *L. chagasi*; D - *L. donovani*; I - *L. infantum*; M - *L. major*; T - *L. tropica*. * WHO reference strains.

Table XLI - Total number of RAPD bands scored per each primer and for all primers for Portuguese *L. infantum*.

	Strain																							
primer	**M***	**A***	**D***	**C***	**I***	**I1**	**I6**	**I8**	**I9**	**I10**	**I11**	**I12**	**I13**	**I18**	**I20**	**I21**	**I22**	**I23**	**I24**	**I25**	**I26**	**I27**	**I29**	**I30**	**total**
A2	4	7	6	5	7	6	6	7	7	6	5	6	8	2	6	6	9	8	8	8	5	5	6	6	17
A4	6	7	8	10	10	10	9	8	9	8	9	9	9	10	10	10	11	11	12	12	12	11	11	11	27
A5	2	4	6	4	5	3	3	6	5	6	4	8	7	6	7	8	7	6	6	8	6	8	6	7	13
A6	12	9	8	8	7	7	9	10	10	10	10	10	10	8	10	8	9	9	8	9	9	8	10	10	23
D3	7	6	7	10	8	6	7	8	8	8	8	8	8	8	7	8	8	8	8	8	8	8	8	8	22
D8	5	0	7	6	5	5	9	10	11	10	9	10	13	8	11	11	11	8	8	11	8	8	10	8	16
D10	8	9	13	14	13	13	14	14	14	14	13	14	14	14	14	14	13	13	14	13	13	13	13	13	24
H1	9	11	16	15	12	13	14	14	15	15	15	15	14	-	15	15	14	14	14	14	13	13	14	14	26
L2	3	3	3	4	3	3	4	5	5	4	4	4	5	5	4	5	5	5	5	4	4	5	5	5	8
total	56	56	74	76	70	66	75	82	84	81	77	84	88	0	84	85	87	82	83	87	78	79	83	82	176

A - *L. aethiopica*; C - *L. chagasi*; D - *L. donovani*; I - *L. infantum*; M - *L. major*. * WHO reference strains.

Table XLII - Total number of RAPD bands scored per primer and all primers in the *L. donovani* complex.

	Primer									
Strain	**A2**	**A4**	**A5**	**A6**	**D8**	**D10**	**H1**	**H4**	**L2**	**total**
M*	6	6	2	3	2	2	5	4	6	36
A*	6	3	4	4	2	3	4	2	5	33
C*	6	10	4	4	4	4	4	4	4	44
I*	7	10	8	6	5	6	5	5	4	56
D*	6	7	7	5	2	4	6	5	4	46
D1	6	9	6	5	3	5	5	5	4	48
D2	5	7	5	5	5	4	6	6	4	47
D3	5	9	4	6	4	4	5	5	4	46
D4	6	5	7	5	4	4	5	5	4	45
D5	6	9	4	5	3	5	6	5	4	47
D6	7	7	5	5	2	4	6	5	4	45
D7	6	7	5	4	4	4	5	5	4	44
D8	-	7	4	-	4	3	4	-	4	26
D9	4	7	5	4	4	-	4	-	5	33
D10	4	8	5	4	5	5	5	3	5	44
D11	3	5	4	2	5	1	5	-	5	30
D12	3	8	5	2	4	5	5	2	5	39
D13	3	8	5	4	5	5	5	3	5	43
D14	4	8	5	4	5	4	5	3	5	43
D15	4	6	5	3	4	3	4	5	5	39
D16	7	7	4	4	4	4	5	6	4	45
D17	3	6	2	4	6	4	2	2	4	33
D18	-	9	2	2	2	4	5	5	5	34
D19	6	6	6	4	3	5	5	4	4	43
D20	4	7	4	2	4	5	5	5	4	40
D21	4	8	3	2	4	4	6	5	4	40
D22	4	8	3	7	3	5	6	5	4	45
D23	4	5	2	5	4	3	3	6	-	32
D24	6	8	2	6	3	5	5	5	4	44
D25	6	8	4	6	4	4	6	6	4	48
D26	6	7	3	5	4	5	3	6	4	43
D27	6	7	4	4	4	4	5	4	4	42
D28	5	7	4	3	5	4	6	6	4	44
D29	6	5	6	3	5	5	5	6	4	45
D30	6	8	4	4	4	4	5	7	4	46
D31	-	10	3	5	4	4	6	6	4	42
D32	7	8	6	3	4	5	5	6	4	48
D33	7	6	5	4	4	5	6	6	4	47
D34	7	6	3	4	4	5	6	5	4	44
D35	7	6	3	3	4	5	6	5	4	43
I31	7	6	4	4	4	5	5	5	4	44
I4	7	9	6	4	4	4	4	4	4	46
I16	7	8	4	4	4	4	5	5	4	45
I17	6	8	4	3	4	5	5	5	4	44
I25	6	9	7	4	4	5	5	5	4	49
I26	6	9	7	-	4	5	5	5	4	45
total	16	24	17	13	14	12	13	13	10	132

A - *L. aethiopica*; C - *L. chagasi*; D - *L. donovani*; I - *L. infantum*; M - *L. major*. * WHO reference strains.

Table XLIII - Bands scored from ITS restriction. Scored 127 total band positions.

Enzyme	*Alu*I								*Bst*UI																							
bp	372	396	404	412	571	653	698	749	65	115	118	123	130	132	134	137	144	147	154	177	186	193	280	290	300	327	360	402	515	530
Strain																														
M*	0	0	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0
A*	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	0	1	1	0	0
T*	1	1	0	0	1	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	1	0	0	0	0	1	1
C*	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
C2	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
C16	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
C13	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
C17	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
C20	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
I*	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
I3	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
I4	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
I16	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
I17	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
I25	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
I26	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
I31	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
I33	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D*	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D1	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D2	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D3	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D4	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D5	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D6	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D7	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D8	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D9	0	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D10	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D11	0	1	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D12	0	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D13	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D14	0	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D15	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D16	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D17	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D18	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D19	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D20	0	1	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D21	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D22	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D24	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D25	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D26	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D27	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D28	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D29	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D30	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D31	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D32	0	1	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0
D33	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0
D34	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0
D35	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0

(see page 287)

(continued table)

Enzyme	*Cfo*I																					*Eco*RI								
bp	57	67	89	90	92	112	136	143	148	152	170	173	182	223	262	311	324	336	367	380	388	201	268	273	283	304	323	445	453	765
Strain																														
M*	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
A*	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
T*	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	1
C*	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
C2	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
C16	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
C13	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
C17	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
C20	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I*	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I3	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I4	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I16	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I17	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I25	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I26	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I31	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
I33	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D*	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D2	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D3	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D4	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D5	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D6	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D7	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D8	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D9	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0
D10	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0
D11	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0
D12	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0
D13	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0
D14	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0
D15	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D16	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D17	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D18	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D19	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D20	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0
D21	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D22	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D24	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D25	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D26	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0
D27	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D28	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D29	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0
D30	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D31	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D32	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0
D33	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D34	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0
D35	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0

(see page 287)

(continued table)

Enzyme	*Eco*RI		*Hae*III																								*Mse*I				
bp	795	830	50	52	63	75	80	83	86	93	97	100	124	137	154	242	247	252	274	292	338	339	260	063	566	0	51	157	220	223	229
Strain																															
M*	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	1		1	0	0	0	0
A*	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0		1	0	0	0	0
T*	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0		1	1	0	0	0
C*	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
C2	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
C16	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
C13	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
C17	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
C20	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I*	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I3	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I4	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I16	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I17	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I25	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I26	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I31	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
I33	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D*	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D1	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D2	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D3	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D4	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D5	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D6	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D7	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D8	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D9	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0		1	0	1	0	0
D10	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0		1	0	1	1	0
D11	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0		1	0	1	1	0
D12	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0		1	0	1	1	0
D13	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0		1	0	1	1	0
D14	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0		1	0	1	1	0
D15	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D16	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D17	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D18	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D19	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D20	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0		1	0	1	1	0
D21	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0		0	1	0	0	1
D22	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D24	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D25	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D26	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0		1	0	1	0	0
D27	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D28	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D29	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0		1	0	1	1	0
D30	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D31	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D32	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0		1	0	1	1	0
D33	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D34	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1
D35	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0		1	0	0	0	1

(see page 287)

(continued table)

Enzyme	*Mse*I					*Sph*I									*Taq*I															
bp / Strain	267	291	667	726	755	98	198	230	327	337	337	352	631	851	128	136	144	148	150	154	159	162	165	168	211	233	253	288	306	315
M	0	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0
A	1	0	0	1	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0
T	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0
C	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
C2	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
C16	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
C13	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
C17	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
C20	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I3	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I4	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I16	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I17	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I25	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I26	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I31	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
I33	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
D	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D1	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D2	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D3	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D4	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D5	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D6	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D7	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D8	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D9	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1
D10	0	0	1	0	0	0	0	0	0	1	1	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1
D11	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1
D12	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1
D13	0	0	1	0	0	0	0	0	0	1	1	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1
D14	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D15	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D16	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D17	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D18	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D19	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D20	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1
D21	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D22	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D24	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D25	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D26	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1
D27	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D28	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D29	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1
D30	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D31	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
D32	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1
D33	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0
D34	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
D35	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0

(see page 287)

(continued table)

Enzyme	*TaqI*								*AluI*	*BstUI*	*CfoI*	*EcoRI*	*HaeIII*	*MseI*	*SphI*	*TaqI*
bp	325	335	344	357	502	508	657	TOTAL								
Strain																
M*	0	0	0	0	0	0	1	27	2	6	3	2	4	3	3	4
A*	0	0	0	1	0	0	0	33	2	6	3	2	4	3	3	4
T*	0	0	0	0	0	0	0	32	2	6	3	2	4	3	3	4
C*	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
C2	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
C16	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
C13	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
C17	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
C20	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
I*	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
I3	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
I4	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
I16	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
I17	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
I25	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
I26	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
I31	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
I33	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D*	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D1	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D2	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D3	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D4	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D5	1	1	0	0	0	0	0	32	2	6	3	2	4	3	3	4
D6	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D7	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D8	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D9	1	1	1	0	1	1	0	41	2	6	3	2	4	3	3	4
D10	0	1	0	0	1	1	0	45	2	6	3	2	4	3	3	4
D11	1	1	1	0	1	1	0	44	2	6	3	2	4	3	3	4
D12	1	1	0	0	0	0	0	37	2	6	3	2	4	3	3	4
D13	0	1	0	0	1	1	0	44	2	6	3	2	4	3	3	4
D14	1	1	0	0	0	0	0	35	2	6	3	2	4	3	3	4
D15	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D16	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D17	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D18	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D19	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D20	1	1	1	0	1	1	0	45	2	6	3	2	4	3	3	4
D21	1	1	0	0	0	0	0	31	2	6	3	2	4	3	3	4
D22	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D24	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D25	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D26	1	1	1	0	1	1	0	38	2	6	3	2	4	3	3	4
D27	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D28	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D29	1	1	1	0	1	1	0	44	2	6	3	2	4	3	3	4
D30	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D31	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D32	1	1	1	0	1	1	0	44	2	6	3	2	4	3	3	4
D33	1	1	0	0	0	0	0	33	2	6	3	2	4	3	3	4
D34	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4
D35	1	1	0	0	0	0	0	34	2	6	3	2	4	3	3	4

A - *L. aethiopica*; C - *L. chagasi*; D - *L. donovani*; I - *L. infantum*; M - *L. major*; T - *L. tropica*. * WHO reference strains. a and b are, respectively, larger and shorter mini-exon amplification products.

Table XLIV - Bands scored from mini-exon RFLP data within the *L. donovani* complex.

Enzyme	*Bst*UI													*Cfo*I											
bp	53	56	59	64	66	81	86	90	93	112	116	120	137	47	51	58	65	75	78	80	82	88	115	117	124	140
Strain																										
M*	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0
A*	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0
T1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0
C*	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1
C2	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
C3	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
C6	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
C13	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
C17	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
C20	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
I*	1	1	1	0	1	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1
I3	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
I4	1	1	1	0	1	1	0	0	0	0	0	1	0	1	1	1	0	0	1	0	1	1	0	0	0	0
I16	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
I17	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
I25	1	1	1	0	1	1	0	0	0	0	1	1	0	1	1	1	0	0	1	0	1	1	0	0	0	0
D*	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0	0	0
D2	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D3	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D4	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D4a	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D5	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1
D6a	1	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0
D6b	1	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0
D7	1	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0
D8a	1	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0
D8b	1	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0
D9	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	0
D10	1	1	1	0	1	1	1	0	1	0	1	1	0	1	1	1	0	0	1	0	1	1	1	1	1	0
D11	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	0
D12	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	0
D13	1	1	1	0	1	1	0	0	0	0	1	1	0	1	1	1	0	0	1	0	1	1	1	1	1	0
D14	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	0
D15	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D16	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1
D17	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1
D18	1	1	1	0	1	1	1	0	0	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1
D19	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1
D20	1	1	1	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	0	0	0	0
D21	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D24	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	0	0	0	1
D25	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1
D26	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1
D27	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	1	1	1
D28	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	1	1	1
D29	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
D30	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1
D32	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	1	1	0
D33	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	1	1	1
D34	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	1	1	1
D35	1	1	1	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	0	1	0	1

(see next page)

(continuation of table)

Enzyme	*HaeIII*														*MspI*											
bp	55	58	62	75	87	93	102	147	154	166	180	190	200	228	52	59	67	69	72	74	78	82	117	180	267	TOTAL
Strain																										
M*	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	13
A*	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	1	1	0	0	0	10
T1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	12
C*	0	1	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	15
C2	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	13
C3	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	13
C6	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	13
C13	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	13
C17	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	13
C20	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	13
I*	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	15
I3	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	13
I4	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	17
I16	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	15
I17	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	14
I25	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	21
D*	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	15
D1	0	1	1	1	1	1	1	1	0	1	0	0	1	0	0	0	0	1	0	0	1	1	1	0	0	30
D2	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	14
D3	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	14
D4	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	15
D4a	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	14
D5	0	1	0	1	1	1	1	1	0	0	0	0	1	0	0	0	0	1	0	0	1	1	1	0	0	32
D6a	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	19
D6b	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	19
D7	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	19
D8a	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	19
D8b	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	19
D9	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	28
D10	0	1	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	1	0	0	1	1	1	0	0	30
D11	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	28
D12	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	28
D13	0	1	0	1	1	1	1	1	0	0	0	0	1	0	0	0	0	1	0	0	1	1	1	0	0	27
D14	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	28
D15	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	16
D16	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	16
D17	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	29
D18	0	1	0	1	1	1	1	1	0	1	0	0	1	0	0	0	0	1	0	0	1	1	1	0	0	32
D19	0	1	0	1	1	1	1	1	0	1	0	0	1	0	0	0	0	1	0	0	1	1	1	0	0	32
D20	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	20
D21	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	15
D24	0	1	1	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	20
D25	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	16
D26	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	16
D27	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	18
D28	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	20
D29	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	17
D30	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	16
D32	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	18
D33	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	19
D34	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	19
D35	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	23
total* *																										51

A - *L. aethiopica*; C - *L. chagasi*; D - *L. donovani*; I - *L. infantum*; M - *L. major*; T - *L. tropica*. * WHO reference strains. ** - Total band positions. a and b are, respectively, larger and shorter mini-exon amplification products.

Table XLV - Total number of bands scored for mini exon RFLP per strain and per enzyme.

	Enzyme				
Strain	***Bst*UI**	***Cfo*I**	***Hae*III**	***Msp*I**	**TOTAL**
M*	3	3	4	3	13
A*	2	2	2	4	10
T1	3	3	2	4	12
C3	4	4	3	2	15
C6	4	4	3	2	13
C13	4	4	3	2	13
C17	4	4	3	2	13
C20	4	4	3	2	13
I*	5	5	3	2	13
I3	4	4	3	2	13
I4	6	6	3	2	15
I16	4	4	3	4	13
I17	4	4	4	2	17
I25	7	6	4	4	15
D*	4	4	4	3	21
D1	10	7	9	4	23
D2	4	4	3	3	15
D3	4	4	3	3	30
D4	4	4	4	3	14
D4a	4	4	3	3	14
D5	10	11	7	4	15
D6a	6	5	5	3	14
D6b	6	5	5	3	32
D7	6	5	5	3	19
D8a	6	5	5	3	19
D8b	6	5	5	3	19
D9	10	9	5	4	19
D9	9	9	8	4	19
D10	9	9	8	4	28
D11	10	9	5	4	30
D12	10	9	5	4	28
D13	7	9	7	4	28
D14	10	9	5	4	27
D15	4	4	5	3	28
D16	4	5	4	3	16
D17	11	10	4	4	16
D18	9	11	8	4	29
D19	10	10	8	4	32
D20	5	6	5	4	32
D21	4	4	4	3	20
D22	0	7	5	4	15
D24	4	7	5	4	20
D25	4	5	4	3	16
D26	4	5	4	3	16
D27	4	7	4	3	18
D28	4	7	5	4	20
D29	4	4	5	4	17
D30	4	5	4	3	16
D32	4	6	5	3	18
D33	4	7	5	3	19
D34	4	7	5	3	19
D35	6	8	5	4	14

A - *L. aethiopica*; C - *L. chagasi*; D - *L. donovani*; I - *L. infantum*; M - *L. major*; T - *L. tropica*. * WHO reference strains. a and b are, respectively, larger and shorter mini-exon amplification products.

Table XLVI - Bands scored from ITG/CS RFLP profiles. Scored 64 band positions.

Enzyme	Size bp	Strains C	C2	C12	C13	C17	C20	I	I3	I6	I16	I17	I26	I31	I33	D	D1	D2	D3	D4	D5	D7	D8	D9	D10
AluI	87	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1
	145	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	160	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1
	180	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
	190	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	0
	245	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	305	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0
	310	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	1
BstUI	65	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	0	1	1	0	0
	70	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	75	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0
	80	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	1	0	0	0	0
	160	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0
	190	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	200	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0
	245	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	265	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	320	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1
	330	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	0	0
	396	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1
CfoI	60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	65	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	0	1
	70	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	77	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	125	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	140	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0
	145	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	148	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0
	149	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1	1	0	0	0	0
	151	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	1
	160	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	180	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	185	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
	201	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
	320	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	0
	455	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	506	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0
HaeIII	65	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	1	1
	140	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	145	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
	147	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0
	180	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	245	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	250	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
MspI	125	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	280	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	290	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1
	300	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0
	517	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
	875	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
RsaI	98	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	137	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
	147	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	0	0	0
TaqI	95	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	201	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	1	1	0	0
	210	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1
	220	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	225	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0
	230	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	315	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	0
	332	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
	345	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	1
total		40	40	40	40	40	40	40	40	40	40	40	39	38	40	36	38	37	37	36	38	35	36	33	36

(continues next page)

(continuation of table)

Enz.	Size bp	Strains D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D24	D25	D26	D27	D28	D29	D30	D31	D32	D33	D34	D35
***Alu*I**	87	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0	0	0	1
	145	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	160	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0
	180	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	190	0	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	245	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	305	0	1	0	0	1	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1
	310	1	0	1	1	0	0	0	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	0
***Bst*UI**	65	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
	70	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	75	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1
	80	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	1	0	1	0	0	1	1	1
	160	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
	190	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	200	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
	245	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	265	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	320	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0
	330	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1
	396	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
***Cfo*I**	60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	65	1	1	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	1	1	1	0	0	0	0
	70	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	77	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	125	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	140	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
	145	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	148	0	0	0	0	0	0	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0	1	1	1
	149	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	151	0	1	1	0	0	1	0	0	0	0	1	0	0	1	0	1	0	1	1	1	1	0	0	0
	160	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	180	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	185	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	201	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	320	0	0	0	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	1	1	0	0	0	0
	455	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0
	506	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	1
***Hae*III**	65	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	100	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	1
	120	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	0	0
	140	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	145	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	147	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
	180	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	245	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	250	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
***Msp*I**	125	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	280	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0
	290	1	1	1	1	0	0	1	1	1	1	0	1	1	0	1	0	1	1	0	0	1	1	1	1
	300	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1	0	1	1	1
	517	1	1	1	1	0	0	1	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	1	1
	875	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
***Rsa*I**	98	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
	137	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	147	0	0	0	0	1	0	1	1	1	0	0	1	0	1	1	1	0	0	1	1	0	1	1	0
***Taq*I**	95	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	201	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	210	1	0	1	1	0	1	0	1	1	1	0	1	0	1	1	1	0	0	1	1	0	0	1	1
	220	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0
	225	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	1	0	0
	230	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	315	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	1
	332	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	345	1	0	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	1	1	1	1	0	0	0
Total		37	35	36	37	39	38	37	37	37	37	37	38	36	41	38	39	35	37	41	39	33	36	36	36

Table XLVII - Total number of fragments scored per strain and per enzyme for ITG/CS RFLP.

	Strains																						
Enz.	**C**	**C2**	**C12**	**C13**	**C17**	**C20**	**I**	**I3**	**I6**	**I16**	**I17**	**I26**	**I31**	**I33**	**D**	**D1**	**D2**	**D3**	**D4**	**D5**	**D7**	**D8**	**D9**	**D10**
AluI	5	5	5	5	5	5	5	5	5	5	5	5	4	5	4	4	4	4	4	4	3	4	5	5
BstUI	9	9	9	9	9	9	9	9	9	9	9	9	9	9	7	7	8	9	8	8	8	8	6	6
CfoI	11	11	11	11	11	11	11	11	11	11	11	11	10	11	9	10	11	11	9	9	9	9	9	10
HaeIII	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	7	6	5	6	6	6	6	6	6
MspI	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	2	2	4	4	4	4	4	4
RsaI	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1
TaqI	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	3	5	4	4	2	4
Enz.	**D11**	**D12**	**D13**	**D14**	**D15**	**D16**	**D1**	**D18**	**D19**	**D20**	**D21**	**D22**	**D24**	**D25**	**D26**	**D27**	**D28**	**D29**	**D30**	**D31**	**D32**	**D33**	**D34**	**D35**
AluI	6	6	5	6	4	4	4	4	4	6	3	4	4	4	5	3	3	6	4	3	3	3	3	4
BstUI	6	6	6	6	9	10	8	8	8	6	10	8	8	11	7	10	8	6	11	10	6	8	8	8
CfoI	10	10	10	10	11	11	9	9	9	10	11	9	9	11	10	11	9	10	11	11	9	9	9	9
HaeIII	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
MspI	4	4	4	4	3	3	4	4	4	4	3	4	4	3	4	3	4	4	3	3	4	4	4	4
RsaI	1	1	1	1	2	0	2	2	2	1	0	2	1	2	2	2	1	1	2	2	1	2	2	1
TaqI	4	2	4	4	4	4	4	4	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4

Table XLVIII - Bands scored from ITG/L RFLP profiles. Scored 55 band positions.

		Strains																									
Enz.	bp	C	C2	C6	C13	C17	C20	I	I3	I4	I16	I17	I25	I26	I31	I33	D	D1	D2	D3	D4	D5	D6	D7	D8	D9
*Alu*I	60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	170	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	175	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	201	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	250	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Bst*UI	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	87	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	103	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0
	105	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	113	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	120	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	126	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	1	0
	130	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
	180	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	270	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Cfo*I	70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	72	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
	75	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0	0	0	0	1
	78	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1
	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0
	85	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	100	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	110	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	124	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	127	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	1	0
	135	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Hae*III	60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	127	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	150	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	162	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	207	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	220	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	240	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Msp*I	122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	142	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1
	146	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
	150	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1
	154	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	0
	255	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	260	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	265	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
	270	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	280	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	298	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	315	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Rsa*I	115	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	510	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	524	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1
	737	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	768	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Taq*I	90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	95	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Total		31	31	31	31	31	32	31	31	31	31	31	31	31	34	34	34	35	33	33	34	36	35	35	35	36
*Alu*I		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
*Bst*UI		6	6	6	6	6	7	6	6	6	6	6	6	6	6	7	8	9	7	7	8	10	8	8	8	6
*Cfo*I		6	6	6	6	6	5	6	6	6	6	6	6	6	8	7	7	7	7	7	7	7	7	7	7	8
*Hae*III		6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
*Msp*I		4	4	4	4	4	5	4	4	4	4	4	4	4	5	5	4	4	4	4	4	4	5	5	5	7
*Rsa*I		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
*Taq*I		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

(continuation of table)

Enz.	Size bp	Strains D10	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D24	D25	D26	D27	D28	D29	D30	D31	D32	D33	D34	D35
*Alu*I	60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	170	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	175	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	201	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	250	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Bst*UI	80	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	0	1	1
	87	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	103	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	0	1	1	0
	105	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	113	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	120	0	0	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1
	126	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
	130	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	1	0
	180	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	270	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Cfo*I	70	0	0	0	0	0	1	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1	1
	72	0	1	1	0	1	0	0	0	0	0	1	0	0	1	1	1	0	1	1	1	1	1	1	0	0
	75	0	1	1	0	1	1	1	0	0	0	1	1	0	1	1	1	0	1	1	1	1	1	1	0	0
	78	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	0	0
	80	0	0	0	0	0	1	1	1	0	0	0	1	1	1	1	0	1	1	0	1	1	0	1	1	1
	85	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	100	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	110	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	124	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	1
	127	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	135	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Hae*III	60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	127	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	150	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	162	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	207	0	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
	220	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	240	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Msp*I	122	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	142	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	0	1
	146	0	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0
	150	0	1	1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0
	154	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	255	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	260	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	265	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	1	0	0	0
	270	0	0	0	0	0	1	1	1	0	0	0	1	0	0	1	0	1	1	0	0	1	0	0	0	0
	280	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	298	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	315	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
*Rsa*I	115	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	510	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	1	0	1	1	1	1	1	1
	524	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	1	0	0	0	0	0	0
	737	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	0
	768	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1
*Taq*I	90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	95	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Total		32	36	34	32	34	40	40	38	35	34	34	38	36	38	40	37	39	39	35	36	38	35	38	37	35
*Alu*I		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
*Bst*UI		6	6	6	6	6	9	9	10	10	9	6	7	9	10	9	8	9	10	7	7	7	6	9	10	8
*Cfo*I		7	8	8	7	8	8	8	8	6	6	8	8	8	9	8	8	7	9	8	8	8	8	9	8	7
*Hae*III		6	6	6	6	6	7	7	6	6	6	6	7	6	6	7	6	7	6	6	7	7	6	6	6	6
*Msp*I		4	7	5	4	5	7	7	5	4	4	5	7	4	4	7	6	7	5	5	5	7	6	5	4	5
*Rsa*I		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
*Taq*I		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Table XXXII - Group assignment of *L. donovani* complex strains.

Code	WHO code	Zym			*MspC*	RAPD	ITS	ITG/ /CS	RFLP	Group
		LON	MON	ILM						
D2	MHOM/KE/1967/MRC(L)3	-	-	3	A1	A	AB	A	A	
D3	MHOM/KE/1973/MRC74	51	-	1	A2	A	AB	A	A	A
D15	MHOM/ET/1984/Addis 142	-	-	3	-	E	AB	A	A	(kenya)
D16	MHOM/KE/1980/Ndandu 4A	44	-	3	A3	(A)	AB	A^h	A	ILM 3
D21	**IMAR/KE/1962/LRC-L57**	44	37	3	A4	A	AB	A^h	A	
D25	MCAN/KE/0000/D2	45	-	8	A4	A	AB	A^h	A	
D27	MHOM/ET/0000/Ayele 8	56	-	3	A5	A	AB	A^h	A	Kenya/
D30	MHOM/KE/0000/Neal R1	56	-	3	-	A	AB	A^h	A	/Ethiopia
D31	MHOM/KE/1975/Mutinga H9	56	32	3	A1	-	AB	A^h	-	
D	**MHOM/IN/1980/DD8†**	41	2	1	B1	B	AB	B	B	
D4	MHOM/IN/1982/Patna 1†	41	-	1	B	B	AB	B	B	B
D6	MHOM/IN/1977/Chowd X†	-	-	1	B	B	AB	-	-	(india)
D7	MHOM/IN/1979/STL1-79†	-	-	1	B	B	AB	B	B	ILM 1
D8	MHOM/IN/1982/Nandi 1†	41	-	1	-	-	AB	B	B	India
D1	MHOM/ET/1967/HU3(LV9)	46	18	5	C	G(C)	AB	C	C	
D5	MMER/IR/1996/Mesh 17	50	-	6	-	F(B)	C	C	C	C
D17	MHOM/LB/1984/Salti 4	-	-	5	C	U	C	C	C	(sudan)
D18	MHOM/SD/0000/Khartoum	46	18	5	C	-	C	C	C	ILM
D19	MHOM/SD/1985/A22	-	-	7	C	G(C)	C	C	C	5/6/7
D22	MARV/SD/1962/LRC-L64	48	-	7	-	U(C)	C	C	-	
D24	MCAN/IT/1976/Dora	50	-	6	C	F(B)	C	C	C	Sudan
D28	MHOM/ET/1972/Gebre 1*	50	82	6	C	U(C)	C	C	C	
D33	**MHOM/SD/1982/Gilani**	48	30	7	C	C	C	C	C	*L. archi-*
D34	MHOM/PT/1992/IMT 180	-	18	5	C	C	C	C	C	*baldi?*
D35	ISER/PT/1993/IMT 188	-	-	5	C	C	C	C	-	
I	**MHOM/TU/1980/IPT 1**	49	1	9	D	D	D	D	D	
C	MHOM/BR/1974/PP75[a]			9	D	D	D	D	D	D
I4	MHOM/ES/1987/Lombardi	-	-	11	D	D	D	-	-	(infantile)
I16	MHOM/CN/1980/Strain A	-	34	9	D	D	D	D	D	ILM 9
I17	MHOM/CY/1963/L53	-	-	12	D	D	D	D	D	(ILM 11)
I25	IARI/PT/1989/IMT 171	-	24	11	D	D	D	-	-	*L. infan-*
I26	IARI/PT/1989/IMT 172	-	24	11	D	D	D	D	D	*tum*
I31	MHOM/MT/1985/Buck	49	78	11	D	D	D	A	-	
D9	MHOM/SA/1987/VL23	-	-	10	-	-	E^h	E	E	
D10	MHOM/SA/1987/VL29	-	-	10	E	E	E^h	E	E	E
D11	MHOM/SA/1987/VL6	-	-	10	-	E	E^h	E	E	(red sea)
D12	MHOM/ET/0000/Ayele 5	52	-	10	E	E	E	E	E	ILM 10
D13	**MHOM/ET/0000/Hussen**	42	-	10	E	E	E^h	E	E	
D14	MHOM/ET/1982/Bekele	42	-	10	-	E	E	E	E	Ethiopia/
D20	MHOM/SD/1987/UGX-marrow	-	-	10	-	E	E^h	E	E	/Saudi Arabia
D29	MHOM/ET/1984/Addis 164	-	83	10	E	U(A)	E^h	E	E	
D32	MHOM/SA/1981/Jeddah KA	42	31	10	-	C	E^h	E	E	
D26	**MHOM/CN/0000/Wangjie 1**	-	35	16	-	U(A)	E^h	F	(E)	F? (china)
C1	MHOM/PA/1980/WR341	-	-	14	-	-	E?	-	-	-
D23	MCAN/IQ/1981/Sukkar 2	43	-	13	I	U	-	-	-	-

In parenthesis are external affiliations to groups. U is ungrouped and in parenthesis, underlined, are majority consensus groups. * - *L. archibaldi*. In bold are strains suggested here as reference strains for their group. [h] are putative heterozygotes. [a] *L. infantum* (*L. chagasi*) alternative reference strain.